Mathematical Contributions to the Theory of Evolution. VI. Genetic (Reproductive) Selection: Inheritance of Fertility in Man, and of Fecundity in Thoroughbred Racehorses

Karl Pearson, Alice Lee and Leslie Bramley-Moore

Phil. Trans. R. Soc. Lond. A 1899 **192**, 257-330
doi: 10.1098/rsta.1899.0006

VI. *Mathematical Contributions to the Theory of Evolution.—VI. Genetic (Reproductive*) Selection: Inheritance of Fertility in Man, and of Fecundity in Thoroughbred Racehorses.*

By Karl Pearson, *F.R.S.*, Alice Lee, *B.A.*, *B.Sc.*, and Leslie Bramley-Moore.

Received November 14,—Read December 8, 1898.

Contents.

---

* The name Reproductive Selection is retained here, although objection has been taken to it, because it has been used in other memoirs of this series. I owe Genetic Selection to Mr. F. Galton.

Introductory.

I UNDERSTAND by a *factor of evolution* any source of progressive change in the constants—mean values, variabilities, correlations—which suffice to define an organ or character, or the interrelations of a group of organs or characters, at any stage in any form of life. To demonstrate the existence of such a factor we require to show more than the plausibility of its effectiveness, we need that a numerical measure of the changes in the organic constants shall be obtained from actual statistical data. These data must be of sufficient extent to render the numerical determinations large as compared with their probable errors.

In a "Note on Reproductive Selection," published in the 'Roy. Soc. Proc.,' vol. 59, p. 301, I have pointed out that if fertility be inherited or if it be correlated with any inherited character—those who are thoroughly conversant with the theory of correlation will recognise that these two things are not the same—then we have a source of progressive change, a *vera causa* of evolution. I then termed this factor of evolution *Reproductive Selection.* As the term has been objected to, I have adopted *Genetic Selection* as an alternative. I mean by this term the influence of different grades of reproductivity in producing change in the predominant type.

If there be two organs A and B both correlated with fertility, but not necessarily correlated with each other,* then genetic or reproductive selection may ultimately cause the predominance in the population of two groups, in which the organs A and B are widely different from their primitive types—'widely different,' because reproductive selection is a source of *progressive* change. Thus this form of selection can be a source, not only of change, but of differential change. As this differentiation is progressive, it may amount in time to that degree of divergence at which crossing between the two groups begins to be difficult or distasteful. We then reach in genetic or reproductive selection a source of the origin of species.

When I assert that genetic (reproductive) selection is a factor of evolution, I do not intend at present to dogmatise as to the amount it is playing or has played in evolution. I intend to isolate it so far as possible from all other factors, and then measure its intensity numerically. If this be sensible, then the demonstration that it is a factor is complete. How far it may be held in check by other factors—*e.g.*, natural or sexual selection—is a matter for further inquiry. If three forces, F_1, F_2, F_3 hold a system sensibly in equilibrium, then F_1 cannot be asserted to be non-effective because no progressive change is visible; its absence would soon bring to light its effectiveness.

The manner in which genetic (reproductive) selection is to some extent held in check will be clearer when my memoir on the influence of directed selection on

---

* If r_{ab} be the correlation of two organic characters A and B, and C be a third character, there is a considerable range of values of r_{ac} and r_{bc}, for which r_{ab} may be zero (*see* YULE, 'Roy. Soc. Proc.' vol. 60, p. 486).

variation and correlation is published. Meanwhile Mr. FILON and I have shown that even a random selection of one organ alters the whole system of correlated organs.[*] Hence genetic (reproductive) selection indirectly modifies not only organs A and B, but all correlated organs. These modifications must be consistent with the maintenance of stamina, physique and fitness to the environment, if the change is not to be counteracted by natural selection.

So far as man is concerned, I have shown[†] that in the case of civilised man, the selective death-rate—*i.e.*, natural selection—does not appear to counteract reproductive selection. A small element of the population produces the larger part of the following generation. I thus concluded that *if* fertility were inherited, reproductive selection was not only a factor of evolution, but in civilised man a very sensible factor, *i.e.*, an apparently incompletely balanced factor.

In the three years which have intervened since writing the essay just referred to, members of the Department of Applied Mathematics in University College, as well as other friends, have occupied their spare time in the collection of data as to fertility and fecundity in the cases of man and of the thoroughbred racehorse. About 16,000 extracts were made in the case of man, and more than 7000 in the case of thoroughbred racehorses. In the course of the work, which proved far more laborious than we had anticipated, many difficulties and pitfalls appeared. But as a general conclusion it seems certain that: *Both fertility and fecundity are inherited, and probably in the manner prescribed by the Law of Ancestral Heredity.*[‡]

The object of this memoir is to set forth the theory and data by aid of which this conclusion was reached. It will be seen that it completes the establishment of genetic or reproductive selection as a factor of evolution by determining the much disputed point as to whether fertility is or is not inherited.

I. *Theory of Genetic or Reproductive Selection.* By KARL PEARSON, F.R.S.

(1.) While the physical result of fertility in an individual is measurable, the quality of fertility or fecundity in an individual differs from other physical characters in that it does not allow of direct measurements except when the potentiality is exerted and the effects recorded. At present we are not able to measure any series of organs or characters in individuals and so ascertain their fertility or fecundity. At the same time there is little doubt that these characters are functions of the physical and measurable organs and characters of the body. Such organs and characters we have good ground for supposing to be inherited according to the Law

---

[*] "Contributions to the Theory of Evolution.—IV. On the Influence of Random Selection on Variation and Correlation," 'Phil. Trans.,' A, vol. 191, p. 234 *et seq.*

[†] "The Chances of Death and other Studies in Evolution. Reproductive Selection," vol. 1, p. 63.

[‡] *See* 'Roy. Soc. Proc.,' vol. 62, p. 386.

of Ancestral Heredity. It seems therefore worth while to prove the following proposition :

Proposition I.—Any character not itself directly measurable, but a function of physically measurable characters and organs inherited according to the Law of Ancestral Heredity, will itself be inherited according to that law.

Thus if we assume intellectual and emotional characters to be ultimately a result of physical conformation, we may be fairly certain that although we know neither the organs of which they are a function, nor the nature of that function, still they will be inherited according to the same law as that which holds for physically measurable organs.

Let y be the character in a parent, and let it be an unknown function f of the unknown physical organs $x_1, x_2, x_3, \ldots x_m$, or let :

$$y = f(x_1, x_2, x_3, \ldots x_m) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i.)}.$$

Let Δy denote the deviation from the mean value of the character y in some special individual, and Δx the deviation from the mean of any x organ in the same individual. Then if these deviations be small compared with the mean values of the organs considered, we have from (i.) above :

$$\Delta y = a_1 \Delta x_1 + a_2 \Delta x_2 + a_3 \Delta x_3 + \ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii.)},$$

where $a_1, a_2 \ldots$ are constants independent of the individual variations.

Let σ denote a standard deviation, ρ a coefficient of interorganic correlation, S a summation with regard to all individuals with character y dealt with, and let them be n in number. Then :

$$n\sigma_y^2 = \mathrm{S}(\Delta y)^2 = \mathrm{S}(a_1 \Delta x_1 + a_2 \Delta x_2 + a_3 \Delta x_3 + \ldots)^2$$
$$= n(a_1^2\sigma_{x_1}^2 + a_2^2\sigma_{x_2}^2 + a_3\sigma_{x_3}^2 + \ldots + 2a_1a_2\sigma_{x_1}\sigma_{x_2}\rho_{x_1x_2} + 2a_1a_3\sigma_{x_1}\sigma_{x_3}\rho_{x_1x_3} + \ldots);$$

or

$$\sigma_y^2 = \Sigma(a_1^2\sigma_{x_1}^2) + 2\Sigma(a_1a_2\sigma_{x_1}\sigma_{x_2}\rho_{x_1x_2}) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)},$$

where Σ denotes a summation through the group of m organs.

Let y' denote the character in an individual who is the offspring of the individual of character y, and $x'_1, x'_2, x'_3 \ldots$ the corresponding organs. Then, if we do not suppose the nature of the function f to have changed in a single generation, we have :

$$y' = f(x'_1, x'_2, x'_3 \ldots x'_m),$$

and

$$\Delta y' = a_1 \Delta x'_1 + a_2 \Delta x'_2 + a_3 \Delta x'_3 + \ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)},$$

$$\sigma_{y'}^2 = \Sigma(a_1^2\sigma_{x'_1}^2) + 2\Sigma(a_1a_2\sigma_{x'_1}\sigma_{x'_2}\rho_{x'_1x'_2}) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(v.)}.$$

Let r be a coefficient of direct heredity expressing the correlation between parent and offspring, and according to the Law of Ancestral Heredity the same for all organs. Then multiplying (ii.) and (iv.) together and summing we have :

$$n\sigma_y \sigma_{y'} \mathrm{R} = \mathrm{S}(\Delta y \Delta y') = \Sigma(a_1^2 \mathrm{S}(\Delta x_1 \Delta x'_1)) + \Sigma(a_1 a_2 \mathrm{S}(\Delta x_1 \Delta x'_2 + \Delta x_2 \Delta x'_1)),$$

where R is the coefficient of correlation between the characters y and y' in parent and offspring. Now :

$$\mathrm{S}(\Delta x_1 \Delta x'_1) = n\sigma_{x_1} \sigma_{x'_1} r$$

$$\mathrm{S}(\Delta x_1 \Delta x'_2 + \Delta x_2 \Delta x'_1) = n\sigma_{x_1} \sigma_{x'_2} r_{x_1 x'_2} + n\sigma_{x'_1} \sigma_{x_2} r_{x'_1 x_2},$$

where $r_{x_1 x'_2}$ and $r_{x'_1 x_2}$ are what I have elsewhere termed coefficients of cross-heredity. Now if the race be stable or sensibly stable for two generations we shall have for all organs $\sigma_{x'} = \sigma_x$. Hence :

$$\mathrm{S}(\Delta x_1 \Delta x'_1) = n\sigma_{x_1}^2 \times r$$

$$\mathrm{S}(\Delta x_1 \Delta x'_2 + \Delta x_2 \Delta x'_1) = n\sigma_{x_1} \sigma_{x_2} (r_{x_1 x'_2} + r_{x'_1 x_2}) = n\sigma_{x_1} \sigma_{x_2} \times 2r\rho_{x_1 x_2},$$

for it is shown in my memoir on the Law of Ancestral Heredity* that on a probable hypothesis :

$$\tfrac{1}{2}(r_{x_1 x'_2} + r_{x'_1 x_2}) = r \times \rho_{x_1 x_2}.$$

Thus we find on substitution :

$$\sigma_y \sigma_{y'} \mathrm{R} = r(\Sigma(a_1^2 \sigma_{x_1}^2) + 2\Sigma(a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2})).$$

But (iii.) and (iv.) show us that $\sigma_y = \sigma_{y'}$, if there be no sensible changes in a generation. Hence :

$$\sigma_y \sigma_{y'} = (\Sigma(a_1^2 \sigma_{x_1}^2) + 2\Sigma(a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2})),$$

and

$$\mathrm{R} = r.$$

Thus the character which is a function of physical organs is inherited at the same rate as those organs themselves.

As we may not unreasonably consider fertility and fecundity to be functions of physically measurable organs, even if we cannot specify which organs, we may, *à priori*, expect fertility and fecundity to be inherited characters.

(2.) *Proposition II.—To determine the numerical values of the changes in mean variation and correlation if fertility be inherited.*

Let us first define two terms which will be frequently used in the sequel.

(*a*.) The *fertility* of an individual shall be defined as the total number of actual offspring.

---

* 'Roy. Soc. Proc.,' vol. 62, p. 411. The hypothesis yet awaits an experimental verification. The need to use it prevents Proposition I. being self-evident.

(*b.*) The *fecundity* of an individual shall be defined as the ratio of the total number of actual offspring to the total number of offspring which might have come into existence under the circumstances.

These definitions are not intended to give precise statistical measures at this stage of our investigations. They are merely meant to convey a general sense of the words, which will be more precisely limited when they are applied to any given species. Fertility and fecundity, as we have thus defined them, leave out of account individual conditions and definite conditions of period, age and environment, which must be fully stated before numerical measures can be made in any special case. When the words are used in this theoretical section the reader must suppose the phrase, "under definite individual and environmental conditions," to be always inserted.

Let M_1 = the mean fertility of parents of one sex; M'_1 = the mean fertility of parents of one sex weighted with their fertility;* N_1 the number of parents considered in the first case, N'_1 the apparent number dealt with in the second case; let σ_1 and σ'_1 be the standard deviations in the two cases, and let x represent the fertility of an individual parent and z its frequency among N_1 parents. Let S denote summation for N_1 parents. Then, without any assumption as to the type of frequency, $N'_1 = S(\lambda xz) = \lambda M_1 N_1$, where λ is a constant such that λx is the weight of a parent of fertility x. This follows at once, since:

$$N_1 = S(z), \qquad M_1 = S(xz)/S(z).$$

Further,

$$\begin{aligned} M'_1 = S(\lambda x \times xz)/N'_1 &= \frac{S(x^2 z)}{M_1 N_1} \\ &= \frac{S\{(x - M_1)^2 z + 2M_1(xz) - M_1^2 z\}}{M_1 N_1}, \\ &= \frac{N_1 \sigma_1^2 + 2M_1^2 N_1 - M_1^2 N_1}{M_1 N_1}, \end{aligned}$$

by the definition of standard-deviation. Hence, finally:

$$M'_1 = \frac{\sigma_1^2}{M_1} + M_1 \quad \text{. . . . .} \quad \text{(i.)}.$$

Further:

$$\sigma'^2_1 = \frac{S\{\lambda x (x - M'_1)^2 z\}}{N'_1} = \frac{S\{(x - M_1 + M_1)(x - M_1 + M_1 - M'_1)^2 z\}}{M_1 N_1}.$$

Hence, multiplying out, we find after some reductions:

$$\sigma'^2_1 = \sigma_1^2 \left(1 - \frac{\sigma_1^2}{M_1^2}\right) + \frac{S\{(x - M_1)^3 z\}}{M_1 N_1} \quad \text{. . . . .} \quad \text{(ii.)}.$$

At first sight it might seem a comparatively easy matter to avoid weighting parents with their fertility, but practically it is almost impossible. For example, if records

---

* *i.e.*, if f be the fertility of a parent, each parent is repeated λf times, where λ is a constant.

are sought of the fertility of mothers in mankind, the women will appear under their husbands' names, and the labour of ascertaining whether two sisters have been included is enormous, when large numbers are dealt with. But if two or more sisters have been included, their mother has been weighted with her fertility, and when we seek the correlation between mother and daughter, it will be between mothers and daughters when weighted with fertility. But a still more serious difficulty arises from the fact that all records are themselves weighted records; the same number are not married from each family, hence we are more likely to find a member of a large family included than a member of a small. The large families, when we seek a record of two generations, are more likely to appear than small families. Precisely the same difficulty occurs when we are dealing with thoroughbred horses; a mare with large fertility is less likely to have all her offspring colts, or all her progeny sold abroad, some one or more will probably ultimately come to the stud, and thus mares of large fertility are, *à priori*, more likely to contribute to our fecundity correlation cards. We do not get over this difficulty by taking the mother and only one of her offspring. The record is still weighted with fertility. The practical verification of this lies in the experience that the fertility of mothers will always be found to be greater than that of daughters, although the fertility of the community may really be increasing; the weighting, of course, excludes sterility in the generation of mothers, but the mere exclusion of the sterile is far from accounting for the whole difference.

What we actually find from our records are M'_1 and σ'_1, but what we want for the problem of heredity are M_1 and σ_1. Equations (i.) and (ii.) do not suffice to determine these, because we cannot evaluate the third moment $S\{(x - M_1)^3 z\}$. We can hardly, even for a first approximation, assume it zero, for the standard-deviation, and therefore the individual variation is large as compared with the mean in the case of fertility, *i.e.*, the distribution is markedly skew.

Turning to offspring of the same sex as the parents, say: let M_2 be the mean fertility of offspring taking *one* only to one parent for the number N_1 of parents, supposing the parents not weighted with their fertility; let M'_2 be the mean in the same case when the parents are weighted with their fertility; and let M''_2 be the mean of *all* recorded offspring of the second generation. Let σ_2, σ'_2, σ''_2 be the standard deviations in the fertility of the offspring for the same three cases, and r, r', r'' be the corresponding coefficients of correlation between fertility in parent and in offspring. It seems to me that r is the coefficient which actually measures the real inheritance of fertility, but that in any correlation table that we can form we shall get r' or r''.

Let y be the fertility of any individual among the offspring, and x the fertility of the corresponding parent; let λx as before be the weighting of the parent, and $\lambda' x$ the number of offspring included in the record, λ' being supposed a constant.*

---

* I have been unable so far to find any sensible correlation between size of family and number married in man, but the point is worth a more elaborate investigation.

We have at once the following results for the total numbers dealt with in each case:

$$N_1 = S(z), \qquad N'_1 = S(\lambda x z) = \lambda M_1 N_1,$$
$$N''_1 = S(\lambda x \lambda' x z) = \lambda\lambda' S(x^2 z) = \lambda\lambda'(\sigma_1^2 + M_1^2) N_1 \quad . \quad . \quad . \quad . \quad . \quad \text{(iii.)}.$$

Turning to the means:

$$M_2 = S(yz)/N_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)}.$$

$$M'_2 = S(\lambda x y z)/N'_1 = [S\{(x - M_1)(y - M_2)z\} + M_1 M_2 S(z)]/M_1 N_1 = M_2 + r\frac{\sigma_1\sigma_2}{M_1} \quad \text{(v.)}.$$

$$M''_2 = S(\lambda x \lambda' x y z)/N''_1 = M_2 + \frac{2M_1\sigma_1\sigma_2 r}{\sigma_1^2 + M_1^2} + \frac{S\{(x - M_1)^2(y - M_2)z\}}{N_1(\sigma_1^2 + M_1^2)}$$
$$= M_2 + r\frac{\sigma_1\sigma_2}{M_1} + r\frac{\sigma_1\sigma_2}{M_1}\frac{M_1^2 - \sigma_1^2}{\sigma_1^2 + M_1^2} + \frac{S\{(x - M_1)^2(y - M_2)z\}}{N_1(\sigma_1^2 + M_1^2)}$$

after some reductions. Now make use of (ii.) and we have:

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1} + r\frac{\sigma_1\sigma_2}{M_1}\frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2} + \frac{S\left\{(x - M_1)^2\left((y - M_2) - r\frac{\sigma_2}{\sigma_1}(x - M_1)\right)z\right\}}{(1 + \sigma_1^2/M_1^2)M_1^2 N_1} \quad \text{(vi.)}.$$

But for normal correlation the equation to the straight line of regression is:

$$y - M_2 = r\frac{\sigma_2}{\sigma_1}(x - M_1).$$

Hence for such correlation the mean value of $y - M_2$ for parents $x - M_1$ is equal to $r\frac{\sigma_2}{\sigma_1}(x - M_1)$ and the summation term would vanish. For skew correlation, Mr. YULE has shown that the line just given is the line of closest fit to the curve of regression. Hence even in the case of fertility, where the correlation is certainly skew, the summation term must be extremely small, or even zero. It follows, therefore, that we may write:

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(vii.)}.$$

There is still another mean which ought to be found, namely, that of parents, M''_1, when all their recorded offspring have been entered on the correlation table. We have:

$$M''_1 = S(\lambda x \lambda' x x z)/N''_1 = S(x^3 z)/\{N_1(\sigma_1^2 + M_1^2)\},$$

or, after some reductions:

$$M''_1 = M_1 + \frac{\sigma_1^2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(viii.)}.$$

I now proceed to the standard deviations for the three cases, and the additional case (σ''_1) for parents.

$$\sigma_2^2 = \text{S}\{(y - \text{M}_2)^2 z\}/\text{N}'_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ix.)}.$$

$$\begin{aligned}\sigma'^2_2 &= \text{S}\{\lambda x(y - \text{M}'_2)^2 z\}/\text{N}'_1 \\ &= \text{S}\{x(y - \text{M}_2 + \text{M}_2 - \text{M}'_2)^2 z\}/\text{M}_1\text{N}_1 \\ &= \frac{\text{S}\{x(y - \text{M}_2)^2 z\} + 2\text{S}\{x(y - \text{M}_2)z\}(\text{M}_2 - \text{M}'_2) + \text{M}_1(\text{M}_2 - \text{M}'_2)^2\text{N}_1}{\text{M}_1\text{N}_1}\end{aligned}$$

Whence, after some reductions, we find :

$$\sigma'^2_2 = \sigma_2^2\left\{1 + r^2\left(\frac{\sigma'^2_1}{\sigma_1^2} - 1\right)\right\} + \frac{\text{S}\left\{(x - \text{M}_1)\left((y - \text{M}_2)^2 - r^2\frac{\sigma_2^2}{\sigma_1^2}(x - \text{M}_1)^2\right)z\right\}}{\text{M}_1\text{N}_1} \quad . \quad \text{(x.)}.$$

Now for a nearly straight line of regression :

$$y - \text{M}_2 = r\frac{\sigma_2}{\sigma_1}(x - \text{M}_1) + \eta$$

where η is uncorrelated with $x - \text{M}_1$. It follows accordingly that $\text{S}\{(x - \text{M}_1)^2\eta z\}$ and $\text{S}\{(x - \text{M}_1)\eta^2 z\}$ will both vanish, since $\text{S}(\eta)$ for an array and $\text{S}(x - \text{M}_1)$ for the whole correlation surface will be zero. Hence the summation term in (x.) is either absolutely zero or extremely small. We have accordingly :

$$\sigma'^2_2 = \sigma_2^2\left\{1 + r^2\left(\frac{\sigma'^2_1}{\sigma_1^2} - 1\right)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xi.)}.$$

Before we proceed to determine σ''_2 and σ''_1 it seems simplest to find the coefficients of correlation r_1 r' and r''. We have :

$$r = \text{S}\{(x - \text{M}_1)(y - \text{M}_2)z\}/(\text{N}_1\sigma_1\sigma_2) \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xii.)}.$$

To find r' we have :

$$r' = \text{S}\{\lambda xz(x - \text{M}'_1)(y - \text{M}'_2)\}/(\text{N}'_1\sigma'_1\sigma'_2).$$

Now

$$y - \text{M}_2 = \frac{r\sigma_2}{\sigma_1}(x - \text{M}_1) + \eta,$$

where η is sensibly un-correlated with $x - \text{M}_1$. Hence :

$$\text{N}'_1\sigma'_1\sigma'_2 r' = \text{S}\left\{\lambda xz(x - \text{M}'_1)\left(r\frac{\sigma_2}{\sigma_1}(x - \text{M}_1) + \text{M}_2 - \text{M}'_2 + \eta\right)\right\}.$$

Expanding, the summations with η vanish, and

$$M_2 - M'_2 - r\frac{\sigma_2}{\sigma_1}M_1 = -\frac{r\sigma_2}{\sigma_1}\left(\frac{\sigma_1^2}{M_1} + M_1\right) \text{ by (v.)}$$

$$= -r\frac{\sigma_2}{\sigma_1}M'_1 \text{ by (i.)}$$

But

$$\sigma'^2_1 = S\{\lambda xz\,(x - M'_1)^2\}/N'_1,$$

thus:

$$N'_1\sigma'_1\sigma'_2 r' = S\{\lambda xz\,(x - M'_1)\frac{r\sigma_2}{\sigma_1}(x - M'_1)\}$$

$$= \frac{r\sigma_2}{\sigma_1}S\{\lambda xz\,(x - M'_1)^2\}$$

$$= \frac{r\sigma_2}{\sigma_1}N'_1\sigma'^2_1.$$

Thus we deduce:

$$r' = r\frac{\sigma_2}{\sigma_1}\frac{\sigma'_1}{\sigma'_2}$$

or:

$$r'\sigma'_2/\sigma'_1 = r\sigma_2/\sigma_1 \quad \text{. . . . . (xiii.).}$$

This result has the simple interpretation *that while the coefficient of correlation is changed, the coefficient of regression is unchanged by weighting fertility, or by reproductive selection.*

This important conclusion is only an illustration of a very interesting theorem, which has been referred to in another memoir* and will be proved generally in a memoir on directed selection, written but not yet published, *i.e.*, that in a wide range of cases selection, whether random or directed (natural and artificial) changes correlation but not regression.

Before proceeding further a general remark will enable us to considerably simplify the otherwise lengthy algebra. Namely, the relation of M''_1, M''_2, σ''_1, σ''_2, r'' to M'_1, M'_2, σ'_1, σ'_2, r' is precisely the same as that of M'_1, M'_2, σ'_1, σ'_2, r' themselves to M_1, M_2, σ_1, σ_2, r. Consequently an interchange of symbols in results already found will lead us to the remaining formulæ needful.

As an illustration of this, let us verify the result we have found for M''_2. By an interchange in (v.):

$$M''_1 = M'_2 + r'\frac{\sigma'_1\sigma'_2}{M'_1},$$

hence using (v.), (i.) and (xiii.), we find:

$$M''_2 = M_2 + \frac{r\sigma_1\sigma_2}{M_1} + r\sigma_2\frac{\sigma'^2_1}{\sigma_1}\frac{1}{M_1 + \frac{\sigma_1^2}{M_1}}$$

$$= M_2 + \frac{r\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/M_1^2}\right),$$

---

* "Contributions to the Theory of Evolution.—V. On the Reconstruction of Stature," 'Phil. Trans.' A, vol. 192, p. 177.

exactly the result reached by a longer process in (vii.). Similarly (viii.) may be deduced from (i.). Applying this to find r'' we have from (xiii.):

$$r''\sigma''_2/\sigma''_1 = r'\sigma'_2/\sigma'_1, \text{ and therefore } = r\sigma_2/\sigma_1 \quad . \quad . \quad . \quad . \quad \text{(xiv.)},$$

a result which again extends the constancy of the regression coefficient under the action of reproductive selection.

Next from (xi.):

$$\sigma''^2_2 = \sigma'^2_2 \left\{1 + r'^2\left(\frac{\sigma''^2_1}{\sigma'^2_1} - 1\right)\right\}$$

$$= \sigma^2_2 \left\{1 + r^2\left(\frac{\sigma'^2_1}{\sigma^2_1} - 1\right)\right\} + r'^2\sigma'^2_2\left(\frac{\sigma''^2_1}{\sigma'^2_1} - 1\right),$$

or using (xiii.) and rearranging:

$$\sigma''^2_2 = \sigma^2_2\left\{1 + r^2\left(\frac{\sigma''^2_1}{\sigma^2_1} - 1\right)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xv.)}.$$

Again by interchanges in (ii.):

$$\sigma''^2_1 = \sigma'^2_1\left(1 - \frac{\sigma'^2_1}{M'^2_1}\right) + \frac{S\{(x - M'_1)^3 z'\}}{M'_1 N'_1} \quad . \quad . \quad . \quad . \quad . \quad \text{(xvi.)}.$$

Here z' stands for λxz, and we should obtain a fourth moment of the original system of unweighted parents by substitution. But it is practically impossible to obtain a correlation table for such a system. Thus it is better to allow the summation term to stand as it is, where it represents the third moment of a system of parents, weighted for fertility owing to the nature of the record, but not weighted with all their recorded offspring. (xvi.) is then a relation between the standard-deviations of parents weighted solely by forming a record and weighted both by this and by their offspring.

Equations (i.) to (xvi.) contain the chief theoretical relations of our subject,[*] and I shall consider some points with regard to them in the following section.

(3.) (*a*.) If we wish to ascertain whether fertility is inherited, we have to discover whether r is or is not zero. Now by (xiv.) r vanishes with both r' and r'', and accordingly either of these will suffice to answer the problem. Still better, we may ascertain the coefficient of regression, and then whether our statistics weight for progeny or not we shall obtain the same value. If there be no secular change taking place in the population, due to something else than reproductive selection, we should expect, provided the Law of Ancestral Heredity holds for fertility, that the regression will be near ·3 for parent and offspring.[†]

---

[*] Two of these formulæ, (v.) and (xi.), were given, but in a less precisely defined manner, in my "Note on Reproductive Selection" of 1896, 'Roy. Soc. Proc.,' vol. 59, p. 303.

[†] See "Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 397.

(*b.*) If no reproductive selection exists, *i.e.*, if fertility be not inherited, then $r = 0$, and

$$\sigma''_2 = \sigma'_2 = \sigma_2, \quad M_2 = M'_2 = M''_2,$$

or, however we form a record of offspring, the mean value and variability of their fertility ought not to be changed. We shall see later that this is very far from the truth, and that these values are in whole or part sensibly affected by the manner in which the record is formed.

(*c.*) Although there be no reproductive selection, M_1, M'_1, and M''_1 will not all be equal, it is impossible that they should be. Further, σ_1, σ'_1 and σ''_1 need not be equal; their degree of sensible divergence will depend on the nature of the primitive frequency distribution for parents.

(*d.*) If fertility be inherited, or reproductive selection be an actual factor of evolution, then we see, by comparing (v.) with (i.) and (vii.) with (viii.), that the mean fertility of mothers will always be *apparently* greater than the mean fertility of daughters. This follows, since r is always less than unity, and if the race be not subjected to secular evolution, other than that due to reproductive selection, σ_2 cannot differ very widely from σ_1.*

(*e.*) An argument from means, as to whether fertility is inherited or not, is very likely to be misleading. We may choose two groups from the record for comparison, neglecting the fact that their frequency in the record is not necessarily that of their frequency in the general population. Thus, if one person, say, in four were married, a marriage record of the community might exhibit the proper frequency of families of four, but it would not do so of families of one. The sort of fallacious arguments we have to be prepared for are, for example:

(i.) That the fertility of the community is diminishing, because M'_2 is less than M'_1.

(ii.) That the fertility of the community is increasing, because M'_2 might be $> M_1$ or M''_2 be $> M'_1$.

(iii.) That fertility is not inherited, because, owing to natural selection, or other factor of evolution, one or other of these means for offspring is sensibly equal to one or other of these means for parents.

Owing to the extreme difficulty of insuring that the method of extracting the record really gives us definitely M'_2, say, and not M''_2 (or M''_2 in part), I have discarded all use of the mean values in attempting to ascertain whether fertility is inherited. The following result, however, is tempting, and might possibly be made

---

* A difference between σ_1 and σ_2 would mark natural selection, sexual selection, or some other factor of secular evolution at work; of *secular*, not periodic, evolution, as parents and offspring must have reached the same adult stage to have had their fertility measured.

use of in direct experiments on breeding insects, where a record could be kept *ad hoc*. It follows at once from (i.), (v.), (vii.) and (viii.):

$$\frac{M'_2 - M_2}{M'_1 - M_1} = \frac{M''_2 - M'_2}{M''_1 - M'_1} = r\,\frac{\sigma_2}{\sigma_1} = \text{coefficient of regression} \quad . \quad . \quad \text{(xvii.)}.$$

It is the second ratio which, I think, might with profit be experimentally evaluated.

(*f*.) Since the mean fertility of daughters loaded with the fertility of their mothers is the fertility of the next generation, and we see that this is always greater than M_2, if r be not zero, it follows that the inheritance of fertility marks a progressive change. The only means of counteracting its influence would be the reduction of M_2 to or below M_1 by the action of other equally potent factors of evolution. For the existence of such factors in man I shall later give evidence.

(4.) *Proposition III.—To extend the results obtained for fertility to the problem of fecundity.*

While the fecundity of an individual can often, at any rate approximately, be measured, the fertility is not ascertainable. Thus we can ascertain the number of occasions on which a brood mare has gone to the stallion and the number of foals she has produced, but her fertility, the produce she might have had, if she had throughout her whole career had every facility for breeding, is unknown to us. But if we proceed to form tables for the inheritance of fecundity, we are met by precisely the same difficulties as in the case of fertility. The more fertile individuals are *à priori* more likely to appear in the record, and will be likely to be weighted again with their fertility when we come to deal with their offspring.*

Now it is certain that fertility must be correlated with fecundity; or, if x now represents the fecundity and f the fertility, we shall have for the *mean* fertility for a given fecundity x an expression of the form $\lambda_0 + \lambda_1 x$, always supposing the regression to be sensibly linear. But the fertility must vanish with the fecundity, hence $\lambda_0 = 0$, and λ_1 is really the ratio of mean fertility to mean fecundity. Thus we may write for the fertility f

$$f = \lambda_1 x + \zeta,$$

where ζ may vary widely, but it is not correlated with x.

If now all the symbols we have used with regard to fertility in Section (2) be interpreted as referring to fecundity, we must weight with a factor λf instead of a factor λx, or with a factor $\lambda\lambda_1 x + \lambda\zeta$. So long as this factor is linear, absolutely no change can be made in the results, for, ζ being uncorrelated with x, all summations including

---

* In the case of sires especially, if we are dealing with thoroughbred horses, their comparative fewness at each period renders it quite impossible to deal with one offspring of each parent only.

$S(\zeta)$ vanish. Thus all the values given for M'_1, M'_2, σ'_1, and σ'_2 remain the same, if their results be interpreted in the sense of fecundity and not fertility. If ρ be the correlation between fecundity and fertility, and σ_1, σ_3 the standard deviations of these quantities, then $\lambda_1 = \rho\sigma_3/\sigma_1$; but we have seen that it is also the ratio of mean fertility to mean fecundity. It follows accordingly that ρ is the ratio of the coefficient of variation in fecundity to the coefficient of variation in fertility. If we may judge by the cases of man and horse, so far as I know the only cases in which fertility and fecundity have yet been examined, a coefficient of variation in fecundity amounts to about 30 per cent., while one in fertility is something like 50 per cent. Thus the correlation of fertility with fecundity would be about ·6. We should expect it to have a high value, perhaps even a higher value than this. In the case of thoroughbred horses, ρ will be the correlation between fecundity and *apparent* fertility. By direct investigation in the case of 1000 brood mares I find its value to be ·5152.

Passing now to the correlations r, r', r'', I observe that the proof given for fertility is valid with but few modifications, if these be fecundity correlations (see p. 266), for the proof involves no expansion of the factor $(\lambda_1 x + \zeta)^2$. Hence we conclude that the regression coefficient for the inheritance of fecundity will not be modified by the nature of the record or the weighting of individuals with their fertility.

When we come to the last series of constants, M''_1, M''_2, σ''_1, σ''_2, we find that these will be modified, owing to the presence of the square factor $(\lambda_1 x + \zeta)^2$, although ζ is not correlated with x. The term ζ^2 now comes in, and $S(\zeta^2)$ will give the standard-deviation of an array of fertilities corresponding to a given fecundity, *i.e.*, $S(\zeta^2) = \sigma_3^2(1 - \rho^2) \times$ number in the array.*

I find after some reductions that M''_2 and M''_1 are given by

$$M''_2 = M_2 + r\frac{\sigma_1\sigma_2}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma^2_1}{1 + \sigma^2_1/(\rho^2 M^2_1)}\right) \quad . \quad . \quad . \quad . \qquad \text{(xviii.)},$$

$$M''_1 = M_1 + \frac{\sigma^2_1}{M_1}\left(1 + \frac{\sigma'^2_1/\sigma^2_1}{1 + \sigma^2_1/(\rho^2 M^2_1)}\right) \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xix.)},$$

the correlation of fertility and fecundity being now introduced into the results.

Clearly the result (xvii.)

$$\frac{M''_2 - M'_2}{M''_1 - M'_1} = \text{coefficient of regression} \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(xx.)}$$

still remains true.

For the remaining two constants σ''_2 and σ''_1, I find, after some rather long analysis in the second case, which it seems unnecessary to reproduce,†

---

* Should the regression not be linear, $\sigma_3\sqrt{(1 - \rho^2)}$ is the mean of the standard-deviations of the arrays.

† In the course of the work the squared standard-deviation of a fertility array is assumed to be the same for all arrays $= \sigma_3^2(1 - \rho^2)$, and λ_1 is given its value $\rho\sigma_3/\sigma_1$. See, however, the previous footnote.

$$\sigma''^2_2 = \sigma^2_2 \left\{1 + r^2\left(\frac{\sigma''^2_1}{\sigma^2_1} - 1\right)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxi.)},$$

$$\sigma''^2_1 = \gamma\sigma'^2_1\left(1 - \frac{\gamma\sigma'^2_1}{M'^2_1}\right) + \gamma\,\frac{1-\rho^2}{\rho^2}\,\frac{\sigma^4_1}{M^2_1} + \frac{S\,\{(x - M'_1)^2 Z'\}}{N'_1 M'_1} \quad . \quad . \quad . \quad \text{(xxii.)},$$

and γ is the factor $\frac{M^2_1 + \sigma^2_1}{M^2_1 + \sigma^2_1/\rho^2}$, or as we can write it

$$\gamma = M'_1 \Big/ \left(M'_1 + \frac{\sigma^2_1}{M_1}\,\frac{1-\rho^2}{\rho^2}\right).$$

If ρ be unity or near unity, *i.e.*, fecundity very closely correlated with fertility, $\gamma = 1$, the second term vanishes and (xxii.) becomes identical with the corresponding fertility formula (xvi.), just as (xxi.) is already identical with (xv.).

Thus we see that the whole series of fecundity relations are strikingly like those for fertility, except that in certain of them—those for M''_1, M''_2, σ''_1 and σ''_2—the correlation ρ of fertility and fecundity is introduced. If ρ be considerable, all the remarks we have made on the fertility formulæ may, *mutatis mutandis*, be applied to the measurement of fecundity.

(5.) *Proposition IV.—To deduce formulæ for finding the correlation between any grades of kindred from the means of arrays into which the kindred may be grouped.*

This problem is of very great practical importance. In the case of Man, families are so small that there is comparatively small difficulty in forming all the possible pairs of brethren, say, for any family; but when we come to animals or insects where the fertility may be extremely large, it is practically impossible to form a correlation table involving 50,000 to 100,000 entries.* One thoroughbred sire may have 50 to 80 daughters, and thus give us roughly 1200 to 3200 pairs of sisters to be entered in a correlation table. Still higher results occur in the case of aunts and nieces. It may be asked why we do not content ourselves with one or two pairs from each parent; the answer is simple: we have not (*e.g.*, in the case of thoroughbred animals, pedigree moths, &c.) a great number of sires, and the sire with 50 offspring cannot, for accuracy of result, be put on the same footing as the sire with only 2 to 4. Our process is really an indirect weighting of our results.

(A.) *To find the coefficient of correlation between brethren from the means of the arrays.*

Let x be the measure of any character or organ in one brother (sister), and x' that of a second brother (sister): let m be the mean of one set of brothers, and m' of the

---

* Even with the reduction in labour, introduced by this proposition and by the use of mechanical calculators, Mr. LESLIE BRAMLEY-MOORE and I took practically a week, of eight-hour days, to deduce two coefficients of correlation, *after* the means of the arrays had already been found.

second set. Let n be the number of brothers in an array, and therefore $\frac{1}{2}n(n-1)$ the number of pairs of brothers in the array. Let σ and σ' be the standard deviations of the two sets of brothers, and r the coefficient of correlation between brothers for the organ in question. Let S denote a summation with regard to all pairs of brothers in the community, and Σ with regard to all brothers in an array. Let N be the total number of brothers in the community. Then if we selected our pairs of brothers for tabulation at random (*e.g.*, not by seniority or other character), we should find $m' = m$ and $\sigma' = \sigma$. Further, by definition of correlation

$$\mathrm{N}r\sigma\sigma' = \mathrm{S}(x-m)(x'-m') = \mathrm{S}\Sigma(x - \mathrm{M} + \mathrm{M} - m)(x' - \mathrm{M}' + \mathrm{M}' - m'),$$

where M and M′ are the means of the two sets of brothers in any array and are clearly equal.

Further, $\Sigma(x-\mathrm{M}) = \Sigma(x'-\mathrm{M}') = 0$, when summed for an array, and $\Sigma(x-\mathrm{M})(x'-\mathrm{M}') = 0$, for there is no correlation within the array when the deviations are measured from the mean of the array. Hence :

$$\mathrm{N}r\sigma\sigma' = \mathrm{S}\{\tfrac{1}{2}n(n-1)(\mathrm{M}-m)(\mathrm{M}'-m')\},$$

or

$$\mathrm{N}r\sigma^2 = \mathrm{S}\{\tfrac{1}{2}n(n-1)\mathrm{M}^2\} - 2m\mathrm{S}\{\tfrac{1}{2}n(n-1)\mathrm{M}\} + m^2\mathrm{N};$$

but

$$\mathrm{S}\{\tfrac{1}{2}n(n-1)\mathrm{M}\} = \mathrm{N}m.$$

Thus, finally,

$$r = \frac{\mathrm{S}\{\frac{1}{2}n(n-1)\mathrm{M}^2\}/\mathrm{N} - m^2}{\sigma^2} \quad \text{(xxiii.)}.$$

This can be written

$$r = \sigma_a^2/\sigma^2 \quad \text{(xxiv.)}$$

where σ_a is the standard deviation of the arrays concentrated into their means and loaded with their sizes ; σ is the standard deviation of all brethren loaded with the number of times they are counted as brethren ; m is the mean of all the offspring loaded with the number of times they are counted as brethren.

Let σ_0 be the standard deviation of offspring, and ρ the correlation between parent and offspring, then the standard deviation of an array of offspring, if correlation be sensibly linear,* will be $\sigma_0\sqrt{(1-\rho^2)}$. We have, further,

$$m = \mathrm{S}(x) = \mathrm{S}\Sigma(x - \mathrm{M} + \mathrm{M}) = \mathrm{S}\{\tfrac{1}{2}n(n-1)\mathrm{M}\},$$

$$\mathrm{N}\sigma^2 = \mathrm{S}(x-m)^2 = \mathrm{S}\Sigma(x - \mathrm{M} + \mathrm{M} - m)^2 = \mathrm{S}\{\Sigma(x-\mathrm{M})^2 + \tfrac{1}{2}n(n-1)(\mathrm{M}-m)^2\}.$$

But

$$\Sigma(x-\mathrm{M})^2 = \tfrac{1}{2}n(n-1)\sigma_0^2(1-\rho^2).$$

---

* See, however, the first footnote p. 270.

Thus:

$$N\sigma^2 = N\sigma_0^2(1-\rho^2) + N\sigma_a^2,$$

or:

$$\sigma^2 = \sigma_0^2(1-\rho^2) + \sigma_a^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (\text{xxv.}),$$

and r may be written:

$$r = \frac{\sigma_a^2}{\sigma_0^2(1-\rho^2) + \sigma_a^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (\text{xxvi.}).$$

Here σ_a can be found from the arrays, and σ_0 and ρ will in many cases have been previously ascertained.

(B.) *To find the correlation between "uncles" and "nephews" ("aunts" and "nieces") from the means of the corresponding arrays.*

Let n_1 be the number of uncles in an array, n_2 be the number of nephews in the associated array, so that n_1n_2 is the number of pairs of uncles and nephews provided by the associated arrays. Let $N = S(n_1n_2)$ be the total number of pairs of uncles and nephews in the community under consideration. Let x be the measure of the organ or character in the uncle, x' in the nephew. Let M and M′ be the means of two associated arrays of uncles and nephews respectively. Let m and m' be the means of all uncles weighted with their nephews and all nephews weighted with their uncles respectively, and let $\overline{\sigma}$, $\overline{\sigma}'$ be the corresponding standard deviations under the same circumstances; r' the correlation of uncle and nephew. Then:

$$Nr'\overline{\sigma}\overline{\sigma}' = S(x-m)(x'-m') = S\Sigma(x - M + M - m)(x' - M' + M' - m').$$

Now $\Sigma(x-M) = \Sigma(x'-M') = 0$, and within the arrays there is no association of individual uncles with individual nephews, *i.e.*, $\Sigma(x-M)(x'-M') = 0$. Thus:

$$Nr'\overline{\sigma}\overline{\sigma}' = S\{n_1n_2(M-m)(M'-m')\} = S(n_1n_2MM') - Nmm',$$

since

$$m = S(n_1n_2M)/N, \qquad m' = S(n_1n_2M')/N.$$

Thus:

$$r' = \frac{S(n_1n_2MM')/N - mm'}{\overline{\sigma}\overline{\sigma}'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (\text{xxvii.}).$$

If $\overline{\sigma}_a$ and $\overline{\sigma}'_a$ be the standard deviations of the means of the arrays of uncles and nephews and R the correlation of these means, the numerator is clearly $R\overline{\sigma}_a\overline{\sigma}'_a$. Thus:

$$r' = R\frac{\overline{\sigma}_a\overline{\sigma}'_a}{\overline{\sigma}\overline{\sigma}'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (\text{xxviii.}).$$

Here the numerator as a whole or in parts is easily found from the means of the arrays. If $\bar{\sigma}_0$ and $\bar{\sigma}'_0$ be the means of unloaded uncles and nephews, we note that they are arrays owing to common parentage, and hence their array standard deviations* will be $\bar{\sigma}_0\sqrt{1-\rho^2}$ and $\bar{\sigma}'_0\sqrt{1-\rho^2}$, ρ being the standard deviation of parent and offspring. As before we find:

$$\left.\begin{array}{l}\bar{\sigma}^2 = \bar{\sigma}_a^2 + \bar{\sigma}_0^2\,(1-\rho^2)\quad.\quad.\quad.\quad.\quad.\quad.\\ \bar{\sigma}'^2 = \bar{\sigma}'^2_a + \bar{\sigma}'^2_0\,(1-\rho^2)\quad.\quad.\quad.\quad.\quad.\quad.\end{array}\right\}\text{(xxviii.).}$$

If, as will probably be the case, there be no secular change between uncles and nephews, then $\bar{\sigma} = \bar{\sigma}'$, $\bar{\sigma}_a = \bar{\sigma}'_a$, $\bar{\sigma}_0 = \bar{\sigma}'_0$, and accordingly $r' = \mathrm{R}\bar{\sigma}_a^2/\bar{\sigma}^2$; whence, using (xxiv.), we have:

$$r' = r \times \mathrm{R} \times \frac{\bar{\sigma}_a^2}{\sigma_a^2}\,\frac{\sigma^2}{\bar{\sigma}^2}\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad.\quad\text{(xxix.).}$$

If we could assume $\sigma_a = \bar{\sigma}_a$ and $\sigma = \bar{\sigma}$, this result would reduce to the very simple form:

$$r' = r \times \mathrm{R}.$$

Now the assumption $\bar{\sigma}_0 = \sigma_0$ is, I think, legitimate, for the distribution for an unloaded array of nephews or uncles should be sensibly that of an array of brethren. But the equality of σ_a and $\bar{\sigma}_a$, which would now involve that of σ and $\bar{\sigma}$, is a much more doubtful point. σ_a and $\bar{\sigma}_a$ mark indeed quite different systems of loading. Both, it is true, are of the form

$$\mathrm{S}\,(nn'\mathrm{M}^2)\,/\,\mathrm{N} - \{\mathrm{S}\,(nn'\mathrm{M})\,/\,\mathrm{N}\}^2,$$

but in the case of brethren $n' = \frac{1}{2}(n-1)$ or n' has perfect correlation with n, while in the case of uncles and nephews n' is only imperfectly correlated with n. The intensity of this correlation depends upon the correlation between the sizes of arrays of uncles and nephews, a quantity which may be very small, or not, according to the nature of the record. Hence it appears necessary in applying the method to make some attempt to appreciate the value of $\bar{\sigma}_a$ as well as σ_a. If this be done R can be found from (xxix.), if not directly. This value of R is not without importance for the inheritance of characters latent in one or other sex.

We have thus reduced the correlations of individuals to a calculation of the correlation of arrays.

(6.) *Proposition V.—To find a measure of the effect of mingling uncorrelated material with correlated material.*

The importance of this investigation lies in the fact that death, restraint, or other

---

* Or, again, the means of the standard-deviations of the arrays.

circumstances, completely screen, in a certain number of cases, both the potential fertility and the real fecundity of man. Precisely similar circumstances, which will be considered more at length later, hinder our obtaining in horses a true measure of fecundity for all cases. We are thus really dealing with a mixture of correlated and apparently uncorrelated material. In what manner does the influence of this mixture effect our results?

Let a group N consist of $n_1 + n_2 + n_3 + n_4$ pairs of individuals. Of these, in the case of n_1 pairs, both individuals have the true value of the character under investigation recorded; in the case of n_2 pairs, neither have the true value recorded; in the case of n_3 pairs, it is the first individual of the pair which has a true recorded value, and the second an apparent or fictitious value; lastly, in n_4 cases, let the fictitious value be in the first and the real value in the second individual of the pair. Then there will be no correlation between individuals in the groups n_2, n_3, n_4. Let r be the correlation in the group n_1 and R that observed in the whole group of $N = n_1 + n_2 + n_3 + n_4$. Let x be the measure of a character in the first, x' in the second individual. Let M and M′ be the means of the total groups of the two individuals and Σ, Σ' their standard deviations. In group n_1 let the corresponding quantities be m_1, m'_1, σ_1, σ'_1, and a similar notation hold for the other sub-groups. Then $m_1 = m_3$ and $\sigma_1 = \sigma_3$; $m_2 = m_4$ and $\sigma_2 = \sigma_4$; while $m'_1 = m'_4$ and $\sigma'_1 = \sigma'_4$; $m'_2 = m'_3$, $\sigma'_2 = \sigma'_3$.

We have at once:

$$\mathrm{M} = \frac{n_1 m_1 + n_2 m_2 + n_3 m_3 + n_4 m_4}{m_1 + m_2 + n_3 + n_4} = \frac{(n_1 + n_3)\, m_1 + (n_2 + n_4)\, m_2}{n_1 + n_2 + n_3 + n_4},$$

while

$$\mathrm{M}' = \frac{(n_1 + n_4)\, m'_1 + (n_2 + n_3)\, m'_2}{n_1 + n_2 + n_3 + n_4}.$$

Further:

$$(n_1 + n_2 + n_3 + n_4)\, \Sigma\Sigma'\mathrm{R} = \mathrm{S}\,(x - \mathrm{M})\,(x' - \mathrm{M}'),$$

by the usual properties of product moments

$$\begin{aligned} &= n_1\sigma_1\sigma'_1 r + n_1 (m_1 - \mathrm{M})(m'_1 - \mathrm{M}') + n_2 (m_2 - \mathrm{M})(m'_2 - \mathrm{M}') \\ &\qquad + n_3 (m_3 - \mathrm{M})(m'_3 - \mathrm{M}') + n_4 (m_4 - \mathrm{M})(m'_4 - \mathrm{M}') \\ &= n_1\sigma_1\sigma'_1 r + n_1 m_1 m'_1 + n_2 m_2 m'_2 + n_3 m_3 m'_3 + n_4 m_4 m'_4 \\ &\qquad - \mathrm{M}\,(n_1 m'_1 + n_2 m'_2 + n_3 m'_3 + n_4 m'_4) - \mathrm{M}'\,(n_1 m_1 + n_2 m_2 + n_3 m_3 + n_4 m_4) \\ &\qquad + \mathrm{MM}'\,(n_1 + n_2 + n_3 + n_4) \\ &= n_1\sigma_1\sigma'_1 r + n_1 m_1 m'_1 + n_2 m_2 m'_2 + n_3 m_3 m'_3 + n_4 m_4 m'_4 - \mathrm{MM}'\,(n_1 + n_2 + n_3 + n_4). \end{aligned}$$

Substituting the values of M and M′ and using the relations between the m's, we find after some reductions:

$$\mathrm{N}\Sigma\Sigma'\mathrm{R} = n_1\sigma_1\sigma'_1 r + \frac{n_3 n_4 - n_1 n_2}{n_1 + n_2 + n_3 + n_4}(m_1 - m_2)(m'_1 - m'_2) \quad . \quad . \quad \text{(xxx.)}$$

Let $\frac{1}{p}$ of the N first individuals and $\frac{1}{q}$ of the N second individuals have fictitious values, then $\frac{p-1}{p}$ N and $\frac{q-1}{q}$ N will have their true values. If, now, there is no correlation between the fictitious values in the two cases, we have at once :

$$n_1 = \frac{(p-1)(q-1)}{pq}\mathrm{N}, \qquad n_2 = \frac{1}{pq}\mathrm{N}, \qquad n_3 = \frac{p-1}{pq}\mathrm{N}, \qquad n_4 = \frac{q-1}{pq}\mathrm{N}.$$

From this it follows at once that

$$n_3 n_4 = n_1 n_2,$$

or the second term in (xxx.) vanishes. Thus :

$$\mathrm{R} = \frac{n_1}{\mathrm{N}}\frac{\sigma_1\sigma'_1 r}{\Sigma\Sigma'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxx.) } bis.$$

Thus R vanishes with r, and no spurious correlation could arise from the existence of fictitious values distributed at random through the correlation table. This result might, indeed, (as it often is tacitly) be assumed by some, but it seems very desirable to have a definite proof.

It remains to consider Σ and Σ'. We have :

$$\begin{aligned}\mathrm{N}\Sigma^2 &= n_1\sigma_1^2 + n_2\sigma'^2_2 + n_3\sigma_3^2 + n_4\sigma_4^2 \\ &\qquad + n_1(m_1 - \mathrm{M})^2 + n_2(m_2 - \mathrm{M})^2 + n_3(m_3 - \mathrm{M})^2 + n_4(m_4 - \mathrm{M})^2 \\ &= (n_1 + n_3)\sigma_1^2 + (n_2 + n_4)\sigma_2^2 \\ &\qquad + (n_1 + n_3)m_1^2 + (n_2 + n_4)m_2^2 - (n_1 + n_2 + n_3 + n_4)\mathrm{M}^2,\end{aligned}$$

or

$$\begin{aligned}\Sigma^2 &= \frac{n_1 + n_3}{\mathrm{N}}\sigma_1^2 + \frac{n_2 + n_4}{\mathrm{N}}\sigma_2^2 + \frac{n_1 + n_3}{\mathrm{N}}\frac{n_2 + n_4}{\mathrm{N}}(m_1 - m_2)^2 \\ &= \left(1 - \frac{1}{p}\right)\sigma_1^2 + \frac{1}{p}\sigma_2^2 + \left(1 - \frac{1}{p}\right)\frac{1}{p}(m_1 - m_4)^2 \\ &= \sigma_1^2 + \frac{1}{p}(\sigma_2^2 - \sigma_1^2) + \left(1 - \frac{1}{p}\right)\frac{1}{p}(m_1 - m_4)^2 \quad . \quad . \quad . \quad . \quad \text{(xxxi.)}.\end{aligned}$$

Similarly :

$$\Sigma'^2 = \sigma'^2_1 + \frac{1}{q}(\sigma'^2_2 - \sigma'^2_1) + \left(1 - \frac{1}{q}\right)\frac{1}{q}(m'_1 - m'_2)^2 \quad . \quad . \quad \text{(xxxii.)}.$$

Now if the introduction of the fictitious values consisted of anything of the nature of a wrong pairing of certain individuals, we should simply have $\sigma_1 = \sigma_2$, $\sigma'_1 = \sigma'_2$, $m_1 = m_2$, $m'_1 = m'_2$ and, accordingly, $\Sigma = \sigma_1$, and $\Sigma' = \sigma'_1$.

In any case, if the percentage of fictitious values be not large, the second and third terms are of the second order of small quantities, since $\frac{1}{p}$ and $\frac{1}{q}$ are small. The maximum value of the third term cannot be greater than $\frac{1}{4}(m_1 - m_2)^2$, and this will be relatively small in the cases to which we shall apply it.

For example, no great changes are made in σ, when we vary the amount of fictitious cases introduced into our fertility tables. m_1 and m_2 do, however, change. Thus $\sigma_1 = \sigma_2 = 3$ approximately, and the range $m_1 - m_2 = 1{\cdot}2$. Hence:

$$\Sigma^2 = 9 + \tfrac{1}{4}(1{\cdot}2)^2, \text{ at a maximum, } = 9{\cdot}36,$$

or,

$$\Sigma = 3{\cdot}06.$$

Thus in this *extreme* case there is only 2 per cent. change in the value of Σ. In such cases accordingly we may take for rough approximations $\Sigma = \sigma$ and $\Sigma = \sigma'$. This leads us to:

$$R = \frac{n_1}{N} r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{xxxii.}).$$

Or, *the reduction of correlation, due to the introduction of fictitious values, is obtained by using as a factor the ratio of actual correlated pairs of individuals to the total number of pairs tabulated.*

This result will be of considerable service when we come to deal with the fecundity of thoroughbred racehorses.

(7.) *Proposition VI.—To obtain a measure of the spurious correlation apparently existing between two organs, when a mixture is made of heterogeneous materials.*

Let x and x' be measures of the two organs, and let there be N pairs of organs formed by i heterogeneous groups containing $n_1, n_2, n_3 \ldots$ pairs with means $m_1, m'_1, m_2, m'_2, m_3, m'_3 \ldots$, &c., standard deviations $\sigma_1, \sigma'_1, \sigma_2, \sigma'_2, \sigma_3, \sigma'_3 \ldots$, &c., and correlations $r_1, r_2, r_3 \ldots$, &c. Let M_1 M' be the means of the whole heterogeneous community, Σ, Σ' the standard deviations, and R the correlation. Then:

$$R\Sigma\Sigma'N = S(n\sigma\sigma' r) + S\{n(m - M)(m' - M')\},$$

where S denotes a summation with regard to all i groups. Now if there were no correlation at all between the organs in any one of the i groups, R for the heterogeneous mixture would still not be zero so long as the second summation did not vanish. This, then, is a measure of the spurious correlation produced by making a mixture of uncorrelated materials.

Now $S\{n(m - M)(m' - M')\}$, remembering the values of M and M′ may be written:

$$S\left\{\frac{n_p n_q}{N}(m_p - m_q)(m'_p - m'_q)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(xxxiv.)}$$

where the summation S now refers to every possible pair p and q of the r groups.

Now it is very unlikely, unless i be very large and the numbers $n_1, n_2, n_3 \ldots$ be taken at random, that this expression will vanish. Suppose even that the means of our heterogeneous groups were uncorrelated, *i.e.*, $S(m - M)(m' - M') = 0$, it is unlikely that $S\{n(m - M)(m' - M')\}$ will also be zero, when n is taken at random. With a comparatively few groups, with numbers taken at random, it is extremely improbable that the principal axes of the i points loaded with $n_1, n_2, n_3 \ldots$ will exactly coincide with the directions of the axes of x and x'.

We are thus forced to the conclusion that a mixture of heterogeneous groups, each of which exhibits in itself no organic correlation, will exhibit a greater or less amount of correlation. This correlation may properly be called spurious, yet as it is almost impossible to guarantee the absolute homogeneity of any community, our results for correlation are always liable to an error, the amount of which cannot be foretold. To those who persist in looking upon all correlation as cause and effect, the fact that correlation can be produced between two quite uncorrelated characters A and B by taking an artificial mixture of two closely allied races, must come rather as a shock.*

The better to illustrate this, I take some data recently deduced by Miss C. D. FAWCETT. She finds for 806 male skulls, from the Paris Catacombs, the correlation for length and breadth ·0869 ± ·0236, and for 340 female skulls, from the same locality, − ·0424 ± ·0365. The existence of the negative sign and the comparative smallness of the correlation, as compared with the probable errors, might lead us to assert the correlation between the length and breadth of French skulls to be sensibly zero.

If now the two sexes be mixed, the heterogeneous group has for correlation ·1968 ± ·0192, a value which cannot possibly be considered zero. Thus the mixture exhibits a large spurious correlation.

Whether any given mixture increases or reduces the correlation will depend entirely on the signs of the differences of the means of the sub-groups. But the danger of heterogeneity for the problem of correlation will have been made manifest. If the value of R for any mixture, whose components are known, is to be calculated, then we have only to note that:

$$\Sigma^2 = \frac{S(n\sigma^2)}{N} + \frac{S(n_p n_q (m_p - m_q)^2)}{N^2}, \quad \Sigma'^2 = \frac{S(n\sigma'^2)}{N} + \frac{S(n_p n_q (m'_p - m'_q)^2)}{N^2} \quad \text{(xxxv.)}.$$

---

* Thus the mere fact of breeding from *two or three* individuals selected at random can easily produce a correlation between organs in the offspring, which has no existence in the species at large.

II. *On the Inheritance of Fertility in Mankind.* By KARL PEARSON, F.R.S., and ALICE LEE, B.A., B.Sc.

(8.) In commencing an investigation of this kind where the results to be expected were quite unknown to us, but where we had reason to believe that the *apparent* strength of inheritance must be very small, we considered that the first thing to be done was to investigate the largest possible amount of material. Thus the probable errors of our results would be very small and any, however small, correlation between fertility in parent and offspring would be brought to light. Attempts might then be made to strengthen any correlation discovered by removing so far as possible one after another the various factors tending to screen the full effect of the inheritance of fertility.

Such factors are for example:

(*a*.) The age of both husband and wife at the time of marriage. The real fertility may be screened by late marriages of one or both parents. The relation of fertility to age at marriage has been dealt with by several writers, notably by DUNCAN and ANSELL.*

(*b*.) The duration of marriage. The data may be taken from a marriage not yet complete, both parents being still alive. Or from a marriage which is complete one or both parents being dead. In the former or the latter case the marriage may be complete so far as fertility is concerned, *i.e.*, details of offspring may be available till the wife has reached the age of 50 years, which for statistical purposes may be taken as an upper limit to fecundity.

(*c*.) Restriction of fertility during marriage. It has been shown in a paper on Reproductive Selection† that there is evidence of the sensible influence of this factor in man. It tends to give fictitious values to the fertility of the younger, rather than the elder generation, and so obscures the correlation.

We have accordingly two problems before us:

(i.) Supposing these and other factors tending to screen the effects of reproductive selection to exist, can we show that it still produces sensible effects in the case of man, and thus demonstrate that fertility is really inherited?

(ii.) Can we by eliminating these factors so far as possible obtain a lower limit to the coefficient of heredity in the case of fertility, and ascertain whether it approximates in value to what we might expect from the Law of Ancestral Heredity?

The first impression of the reader may be that it is only needful to select the

---

* J. MATHEWS DUNCAN, 'Fecundity, Fertility, Sterility and Allied Topics,' second edition, Edinburgh, 1871. CHARLES ANSELL, Junr., 'Statistics of Families in the Upper and Middle Classes,' London, 1874.

† 'The Chances of Death and other Studies in Evolution,' vol. I, pp. 77, 89.

fertility of marriages, which were formed with husband and wife between 20 and 28* say, and which have lasted till the wife is over 50. But these conditions must be true in *two* successive generations, and, had we adopted them, we may safely say that without immense labour it would have been impossible to collect even a thousand cases. From the whole of the peerage, the baronetage, the landed gentry, a variety of family histories, of private pedigrees, and a collection of data formed of families at first hand, it was not possible to extract more than about 4000 cases for the inheritance of fertility in the female line, when the limitations were far less stringent, being applied only to *one* generation, and consisting in our taking marriages entered into at any time of life for either husband or wife, and lasting till the death of one member or for at least fifteen years. Even in this case the pedigree of the wife had to be sought for from one record to another and often in vain. It is the male pedigree with which the recorder in nearly all cases occupies himself.

Only those who have attempted the labour of extracting, as has been done in this case, some 16,000 separate returns, will fully grasp the difficulty of making the limitations of selection more and more complex; the quantity to be obtained becomes dangerously small and the labour immensely increases. Even could with time and patience a sufficient selection of ideal cases have been made, it does not follow that the result would be satisfactory; for, we should have made a narrow *selection*, and this very fact might indicate that possibly we have been selecting one grade or class of fertility. It is possible that the less fertile are the weaker, and so more liable to die early; or again it may be the more fertile women who are subjected to the more frequent risk of childbed, and thus are less likely to appear in the selection of long marriages. Even greater or less risk at birth may be an inherited character in women, and may not unfairly be looked upon in itself as a factor limiting fertility naturally.

Taking these points into consideration, it seemed that if we were to have enough material to draw conclusions from we must entirely drop all attempt to classify by age of parents at marriage. We might make some limitations but they must not be very stringent; they must leave room for an increase of stringency in different directions, so that we could roughly appreciate the influence of the screening factors. Accordingly our plan has been to show that correlation actually does exist between parent and offspring with regard to fertility, and that when we make the conditions more stringent the correlation increases towards the value indicated by the law of ancestral heredity.

(9.) *On the Inheritance of Fertility in Woman.*—(i.) Table I. gives the result for 4418 cases of the fertility of a mother and of her daughter. These were extracted from FOSTER's 'Peerage and Baronetage,' BURKE's 'Landed Gentry,' some family

---

* As DUNCAN points out, an early marriage on the average means an earlier cessation of fecundity; a somewhat later one does not necessarily connote less fertility.

histories and a collection of family data drawn from private pedigrees and other sources. In the case of the daughter, no marriage was taken which had not lasted at least 15 years, or until the death of husband or wife. In the case of the mother no limitation whatever was made, the number of brothers and sisters of the daughter, including herself, being counted. Weight was given to the fertility of the mother, for every possible case that could be got from the records under the above conditions was extracted. It is quite possible that a certain proportion of offspring dying in early infancy have not been entered in the records.

If M_m, M_d be the mean fertilities of mother and daughter, σ_m, σ_d their standard deviations, and r_{md} their correlation, we found :

$$M_d = 3{\cdot}494, \qquad M_m = 6{\cdot}225,$$
$$\sigma_d = 2{\cdot}975, \qquad \sigma_m = 3{\cdot}052,$$
$$r_{md} = {\cdot}0418.$$

Clearly owing to the near equality of σ_d and σ_m the regression of daughter's on mother's fertility is sensibly equal to the correlation.

The probable error of r_{md} is determined by the formula given by PEARSON and FILON* to be ·0101, or r_{md} is four times its probable error.

We thus conclude :

(i.) That fertility is inherited in the female line.

(ii.) That its effects are very largely screened by the factors to which we have previously referred.

Had we started with no limitation as to the daughter's family, it is highly probable that r_{md} would scarcely have been sensible relatively to its probable error, and, therefore, small series without due regard to screening causes may easily lead the recorder to suppose that fertility is not inherited.

Supposing we exclude from the daughters the 775 barren marriages, we find the mean for 3643 cases of fertile marriages to be 5·237. Comparing this fertility with the observed fertility 6·225 of mothers, a superficial inquirer might at once consider that a diminution of fertility has taken place. The fact is that neither of the results, M_m or M_d gives the actual fertility of the mothers or daughters. These are the means M''_1 and M''_2 of formulæ (viii.) and (vii.) of the theoretical investigation.

Let us apply the theory developed to our statistics. In the first place we note that r is small ; hence r^2 is still smaller, and thus by (xv.) σ''_2 will not differ much from σ_2. Since σ''_1 will be generally less than σ_1 by (xvi.), it follows that σ''_2 will probably be less than σ_2. Approximately, we can take $\sigma_2 = 3$. Turning to (vii.) we see that M''_2 cannot, since r is small, differ widely from M_2. If there be no secular

---

* "Contributions to Theory of Evolution.—IV." 'Phil. Trans.,' A, vol. 191, p. 242.

evolution in the real fertility sensible in the one generation, then M_1 would equal M_2. Hence to a first approximation we should have :

$$M_1 = M_2 = M''_2 = 3{\cdot}494.$$

To obtain a second approximation we may substitute this in the small terms of (vii.). Here σ'^2_1 must be found from (ii.) ; neglecting the cubic term we have :

$$\sigma'^2_1/\sigma^2_1 = 1 - \sigma^2_1/M^2_1 = {\cdot}2628.$$

Hence :

$$M_1 = M_2 = 3{\cdot}494 - r\frac{\sigma^2_1}{M_1}\left(1 + \frac{{\cdot}2628}{1{\cdot}7372}\right)$$

$$= 3{\cdot}494 - {\cdot}0418 \times 2{\cdot}5759 \times 1{\cdot}1513$$

$$= 3{\cdot}494 - {\cdot}124 = 3{\cdot}370.$$

We can now substitute this value of M_1 in (viii.), and we find :

$$M''_1 = 3{\cdot}370 + 2{\cdot}980 = 6{\cdot}350.$$

This differs comparatively little from the actually observed value, 6·225, and is satisfactory evidence of the validity of our theory. The fact that the elder generation was in no way limited like the younger, and that we have neglected the third moment—although fertility distributions are never normal—as well as made other approximations, is quite sufficient to account for the difference observed.

We may take it that 3·4 is practically the fertility of the elder generation, and that this is raised to about 3·5 by reproductive selection in the younger generation. The result 6·2 for the elder generation is thus purely a result of weighting due to the nature of the record.

(ii.) Table II. gives the result of 1000 cases taken from the Peerage. Here the conditions of extraction were as follows :—

One member only was taken out of each family, or no weight was given to the fertility of mothers.

The daughters' marriages had all been completed by the death of one parent or had lasted at least 15 years.

There was no limitation with regard to the parents' marriages. We found :

$$M_d = 3{\cdot}923, \qquad M_m = 5{\cdot}856,$$

$$\sigma_d = 2{\cdot}758, \qquad \sigma_m = 2{\cdot}751,$$

$$r_{dm} = {\cdot}2096.$$

The coefficient of regression is sensibly equal to that of correlation. The probable error of $r_{dm} = {\cdot}0204$, or not a tenth of the value of r_{dm} itself. Again we conclude that fertility is certainly inherited in the female line. By selecting fairly homogeneous material with a more definite and complete record than exists for the heterogeneous material of the previous case, we have carried up the correlation to five times its previous value, and within a reasonable distance of the value ·3 which would be required by the law of ancestral heredity. The homogeneity of our material is evidenced by the reduction in both standard deviations; the greater completeness of the record by the rise in the fertility of daughters; and the non-weighting of the fertility of mothers by the fall in their mean fertility.

If the reader will turn back to the theory of the influence of heterogeneity on correlation in section (7), he will notice that the expression in (xxxiv.) will be negative, and therefore the apparent correlation less than the real, if we form a mixture of two groups in which $m_p > m_q$ and $m'_p < m'_q$. Now the entries of women in the Landed Gentry and other records are very often entries of "heiresses," while the entries of women in the Peerage are entries because of class. An "heiress" naturally has fewer brothers and sisters than another woman on an average, or we may expect $m'_q > m'_p$. On the other hand an "heiress" need not have fewer children than other women, unless her heritage is the result of her coming from an infertile stock, and is not a result of the incompleteness of her parents' marriage. If she belongs to a somewhat lower social grade, she may possibly be more fertile than the average of a higher social grade. In this case m_p will be $> m_q$, and when we come to mix records of the Peerage with those of the Landed Gentry and Family Histories, we need not be surprised to find the correlation of fertility much weakened, as it undoubtedly is (as shown by (i.) and (ii.) above) by the mixture.

Let us next apply our theory to the above results. We are now dealing with M'_1, M'_2, σ'_1, σ'_2. Assuming that there is no secular change $\sigma_1 = \sigma_2$, and accordingly since $\sigma'_1 = \sigma'_2$ sensibly, formula (xi.) shows us that both $= \sigma_1$.

Further, if $M_1 = M_2$, formula (v.) is a quadratic equation to find M_1; substituting for M'_1, σ_1^2, and r, we have, on solving and taking the only admissible root, $M_1 = 3·4625$. Then, applying formula (i.) to find M'_1, we have:

$$M'_1 = 5·660.$$

This is not quite as high as the observed value 5·856, but it suffices to show that our theory expresses the main facts. In all probability we have not entirely freed our results from weighting with fertility; because, although every endeavour was made to take only one from each family, it is possible that pairs of sisters have occasionally crept into the record.

(iii.) Table III. gives the result of 1000 cases taken from the Landed Gentry. As we have already noted, the women recorded are largely "heiresses," and we believed this might be one of the chief sources of the heterogeneity of the material in Table I. The conditions of selection were made somewhat more stringent, and were as follows:—Only one daughter was taken from each family, and her marriage must have lasted at least 15 years. No limitation was placed on the duration of the parents' marriage. We found :

$$M_d = 4{\cdot}232, \qquad M_m = 5{\cdot}403,$$
$$\sigma_d = 3{\cdot}292, \qquad \sigma_m = 3{\cdot}241,$$
$$r_{md} = {\cdot}1045.$$

The probable error of $r_{md} = {\cdot}0211$, and again we see that fertility is certainly inherited. The correlation has, however, sunk ; probably, as the great increase of variation indicates, because we are dealing with much more heterogeneous material than in the case of the Peerage. While the selection of "heiresses" has largely reduced the number of brothers and sisters, *i.e.*, the fertility of mothers, the limitation to marriages of at least 15 years has increased the apparent fertility of daughters ; nor is this increase at all balanced by the fact that heiresses come from small families, and may, therefore, be supposed to be the children of rather sterile mothers. The average number of children of heiresses is sensibly as large as the average number of children of women who are not in the bulk heiresses, and who have, as in the following case, been selected with the same condition as to duration of marriage. The fact is that heiresses are not on the whole the children of sterile mothers ; their high fertility and *their small correlation with their mothers* shows us that heiresses in the bulk are rather the daughters of mothers whose apparent fertility is fictitious. They have, owing to the sterility or early death of their husband, to their own marriage late in life, or to some physical disability, or other restraint, never reached their true fertility. If this conclusion be correct, and a comparison of the values of M_d and r_{md} in this and the following cases thrusts it almost irresistibly upon us, then we see that the argument against the inheritance of fertility based upon the fertility of heiresses and non-heiresses is of no validity.* It could not be valid as against the values of the correlation we have found, but the present investigation shows by the value of r_{md} exactly wherein the error lies : the heiress is not infertile, but is the daughter of a fictitiously infertile mother.

Applying our theory to this case, we find from formula (xi.), putting $\sigma_1 = \sigma_2$:

$$\sigma_1^2 = (\sigma'^2_2 - r^2\sigma'^2_1)/(1 - r^2),$$

whence we find $\sigma_1 = 3{\cdot}293$, a result sensibly identical with σ_d. Solving the quadratic (v.) with $M_1 = M_2$ to find M_1 we find :

$$M_1 = 3{\cdot}952.$$

Hence by (i.) we have :

$$M'_1 = 6{\cdot}838,$$

the actually observed value being 5·403. Thus the theory completely fails to give the fertility of the heiresses' mothers ; for such a fertility as we find in the daughters,

---

* See, for example, a recent letter of Mr. HOWARD COLLINS in 'Nature,' November 3, 1898.

the mothers' fertility is far too low. This again emphasises the point we have already referred to. The peculiar character of the selection, which leads to the female record in the Landed Gentry, is not one such as we have considered in our theory, where the record of any family is likely to appear in proportion to its size. Such a distribution is a *chance* distribution, but a selection of women inheriting land has not this character, and a woman who is the mother of co-heiresses is hardly doubly as likely to appear as the woman who is mother of one. A marriage in either case is likely to be arranged, and if we take only one daughter from each family the record will not already have weighted—at any rate to the full extent—every mother with her fertility. If the reader will compare the variation columns for both daughters and mothers in Table III. with the corresponding columns in Table II. or Table IV., he will at once see how anomalous is the selection of women given in the Landed Gentry.

(iv.) Table IV. gives the results for 1000 cases taken from the Peerage and Baronetage under the following limitations: one daughter only was taken for each mother, and in the case of both mother and daughter the marriage must have lasted at least 15 years. We found:

$$\mathrm{M}_d = 4{\cdot}335, \qquad \mathrm{M}_m = 5{\cdot}898,$$

$$\sigma_d = 2{\cdot}967, \qquad \sigma_m' = 2{\cdot}830,$$

$$r_{md} = {\cdot}2130.$$

The probable error of $r_{md} = {\cdot}0204$. Thus, as it is now hardly necessary to repeat, fertility is certainly and markedly inherited. The regression coefficient is now as high as ·2233, the closest limit we have yet reached to the theoretical ·3 of the law of ancestral heredity.

Owing to the limitation to marriages of 15 or more years, the means of the fertilities of both mothers and daughters have risen, in the latter case more, however, than the former. It might have been expected that the fertility of mothers would have risen more, but it must be remembered that M_m is the apparent and not the *real* fertility of mothers; and further, since the record largely weights the more fertile women, the bulk of the mothers are already those with large families, *i.e.*, those whose marriages have lasted at least 15 years.

Assuming that there is no sensible secular change in unweighted fertility, *i.e.*, $\sigma_1 = \sigma_2$, we have from the formula on p. 284:

$$\sigma_1 = 2{\cdot}973.$$

From (v.) with $\mathrm{M}_1 = \mathrm{M}_2$ we find:

$$\mathrm{M}_1 = 3{\cdot}845$$

for the real fertility of mothers. This is a sensible increase on the value 3·463 given in Case (ii.), in which there was no minimum duration to the length of the mother's marriage.

Applying formula (i.) we find:

$$M'_1 = 6{\cdot}144,$$

which is somewhat more than the observed value 5·898. The reason for this lies, we think, in the difficulty already referred to on p. 263. If we start extracting mothers, it is often difficult to follow the daughter's history; starting with the daughter it is much easier, although still laborious, to trace back her ancestry, and find the number of her brothers and sisters. Even in this case the search may be lengthy. But as daughters when married change their name, it requires great care in extracting large quantities to be sure that a mother is not repeated, *i.e.*, some approach made to weighting her with her fertility. Every care was made in extracting the records, but we cannot hope to have always avoided weighting to some extent a mother, and if this be done we shall have a transition from formulæ (xi.), (v.), and (i.) towards formulæ (xv.), (vii.), and (viii.), which would well account for the difference found between theory and observation.

If we sum up for inheritance of fertility in the female line on the basis of these four cases, we draw from each one of them the unquestionable result that fertility in woman is an inherited character. Further, the more we remove causes of fictitious values for the fertility in either generation, the closer does the value approach that required by the law of ancestral heredity. The two chief disturbing factors which we have not been able to eliminate are (*a.*) the age at which marriage is entered upon, (*b.*) restraint giving a fictitious value to the fertility. Both these causes must give a lessened value to the correlation of fertility between mother and daughter, and the first, judging from the great influence of age at marriage on fertility, cannot fail to give a serious diminution. Hence if we find the regression coefficient as high as ·2233, when we neglect these factors, it is no stretching of facts to conclude that it would in all probability rise to ·3 could we take them into account.

Our conclusion, therefore, is that fertility in woman is certainly inherited through the female line, and most probably according to the law of ancestral heredity. Reproductive selection is actually a *vera causa* of progressive change, but its influence is largely, if not entirely screened by the numerous factors tending to make the apparent fertility of women differ from their real or potential fertility.

(10.) *On the Inheritance of Fertility in Man.*

(i.) While many of the difficulties involved in the extraction of data for women still exist for man, a new and important feature tending to screen the full influence of the law of ancestral heredity arises in his case. The full fertility of the husband is not in the average case at all approached in the case of monogamic marriage. Hence, in considering the size of a man's family as a measure of his fertility we are measuring a character which differs largely from the character of fertility in woman. It is only in the case of sterile or even very sterile men that there is likely to be a correlation shown between the sizes of the families of fathers and sons.

The intensity or duration of fecundity in the husband must, one or other, be less than that of the wife,—and this will hardly be so in the great run of cases—if his family is to be in any way a measure of his fertility, or, as it might be better to call it in this case, his sterility. We are seeking to find a correlation between two characters, one in father and one in the son, neither of which we can measure unless they fall short of a certain limit. The result is that our correlated material is weakened down by the admixture of a mass of uncorrelated material in the manner indicated in Proposition V. of the theoretical part of this investigation. Within the family we cannot hope to get a correlation which will approach that indicated by the law of ancestral heredity. We may still, however, hope to ascertain whether fertility, respectively sterility, is an inherited character in man as well as woman.

(ii.) Our first attempt was to collect as much material as possible, so that our limitations were few. The Peerage, Baronetage, Landed Gentry, Family Histories, private pedigrees, and collected data provided the 6,070 cases arranged in Table V. Here large families were weighted because several, where available, were taken from one family. The son's marriage must either have lasted till the death of one partner or at least 15 years; there was no condition as to the duration of the father's marriage.

We have spoken of the correlation between fertility of father and son, but since only a single marriage of the father is taken, it may be equally well termed a correlation between the fertility of the mother and son, which may, perhaps, to some extent explain the relatively high values reached.

Let M_s, σ_s be the mean and standard-deviation of the son, M_p, σ_p of the parent, and r_{sp} the correlation; then we found:

$$M_s = 3{\cdot}871, \qquad M_p = 5{\cdot}831,$$
$$\sigma_s = 3{\cdot}003, \qquad \sigma_p = 3{\cdot}190,$$
$$r_{sp} = {\cdot}0514.$$

The probable error of $r_{sp} = {\cdot}0087$. Thus the correlation is nearly six times the probable error, or fertility in man is certainly inherited.

(iii.) Table VI. contains the result of extracting 1,000 cases from the Peerage, only one son being taken from each family, and his marriage having lasted at least 15 years. No attention was paid to the length of parents' marriage.

We found:

$$M_s = 5{\cdot}070, \qquad M_p = 5{\cdot}827,$$
$$\sigma_s = 2{\cdot}910, \qquad \sigma_p = 3{\cdot}142,$$
$$r_{sp} = {\cdot}0656.$$

The probable error of $r_{sp} = {\cdot}0212$. This case closely confirms the previous case; M_p and σ_p remain sensibly the same, M_s has risen owing to the longer period of duration of the son's marriage, and since there is a longer period for the possible exhaustion of the male fertility, we find r_{sp} is slightly larger. Although the numbers are smaller than in Case (i.), the probable error is not so large but that we can still assert an inheritance of fertility in man.

(iv.) Lastly, to compare with Case (iii.) for women, 1000 cases were extracted from the Landed Gentry, and are given in Table VII. Here no marriage of the son or parents was taken under a minimum of 15 years' duration, and only one son taken from each family. We found :

$$M_s = 5{\cdot}304, \qquad M_p = 6{\cdot}272,$$
$$\sigma_s = 2{\cdot}951, \qquad \sigma_p = 2{\cdot}911,$$
$$r_{sp} = {\cdot}1161.$$

Thus the longer duration of the marriage, which gives a greater chance for the exhaustion of the fertility of a partially sterile father, leads to an increased correlation. The probable error here is ·0210, and the correlation is thus unquestionable.

It would be idle to apply the theory before developed to these male cases, for the simple reason that we must certainly look upon them as containing a large proportion of uncorrelated material. But they suffice to show that male fertility is an inherited character, and although the results are widely different from those indicated by the law of ancestral heredity, they are large when we consider how little male fertility appears measurable by the results of monogamic marriage. Were an approximately close measure of male fertility available, there is certainly in the above results no reason to induce us to believe that it would not be found to obey the law of ancestral heredity.

(11.) *On the Inheritance of Fertility in Woman through the Male Line.*

Although we are not able to measure the potential fertility of the male, we are able to determine whether he transfers fertility from his mother to his daughter. This may be simply done by correlating the fertility of a woman and that of her paternal grandmother. This problem belongs to an important class—namely, questions as to the extent to which a sexual character is inherited through the opposite sex. DARWIN has touched upon this "transmission without development" in Chapter viii. of the 'Descent of Man,'* and we shall find his views amply verified.

The problem before us is : Does a woman have as close correlation with her paternal as with her maternal grandmother in the matter of fertility ?

To solve this problem 1000 cases were taken out of the Peerage for the fertility of a woman and of her paternal grandmother. The marriages of the woman and of her grandmother were both taken with a minimum duration of fifteen years. Every care was taken that no weight should be given to fertile families by taking only one out

---

* Second Edition, p. 227, 'Laws of Inheritance.'

of each family, but, of course, the difficulty of avoiding this is increased when a pedigree must be traced through three instead of two generations.

If d denotes granddaughter, g grandmother, the following results were obtained (Table VIII.) :—

$$\mathrm{M}_d = 4\cdot411, \qquad \mathrm{M}_g = 5\cdot657,$$
$$\sigma_d = 2\cdot897, \qquad \sigma_g = 3\cdot056,$$
$$r_{dg} = \cdot1123.$$

The coefficient of regression of daughter's fertility on grandmother's fertility $= \cdot1065$.

The probable error of $r_{dg} = \cdot0211$.

According to the law of ancestral heredity* we should expect the grandparental correlation and regression to be half the parental and equal to ·15. Comparing the present result with Case (iv.), we see that ·1123 and ·1065 have to be compared with $\frac{1}{2}$ (·2130) and $\frac{1}{2}$ (·2233), or with ·1065 and ·1116. These are differences well within the probable error of our results, or we may conclude that the correlation of a woman with her paternal grandmother is exactly what from Case (iv.) of Section (9) we should expect to find for her correlation with her maternal grandmother. The reduction from ·15 to ·1123 is just what we might have predicted after the maternal reduction from ·3 to ·2130. We, therefore, conclude that the fertility of woman is inherited through the male line with the same intensity as through the female, and this intensity is most probably that which would be indicated by the law of ancestral heredity.

(12.) We do not stay to consider many points which flow from our tables, such, for example, as the amount of restraint indicated by the hump at the start of our various frequency distributions for size of families, partly because such consideration would lead us beyond our present scope, the inheritance of fertility, and partly because this point has been already dealt with by one of us in a paper on 'Reproductive Selection.' We consider that we have shown fertility in mankind to be an inherited character in both lines, and probably obeying the law of ancestral heredity.† By aid of our theoretical investigations it is clear that the average size of a family (M_1), as deduced from our record data (M'_1 or M''_1), is about 3·5 children, if the marriage lasts till the death of one partner, or at least till 15 years; it is about 3·9 to 4 children if the duration of the marriage is at least 15 years. Reproductive selection would increase this average by about ·5 child per generation were its influence not counteracted

---

* "Mathematical Contributions to the Theory of Heredity, on the Law of Ancestral Heredity," 'Roy. Soc. Proc.,' vol. 62, p. 397.

† In the paper on "The Law of Ancestral Heredity" ('Roy. Soc. Proc.,' vol. 62, p. 412) it is stated that fertility is probably inherited, but the amount falls below that which would be indicated by the law of ancestral heredity. At that time only Case (i.) of Section (9) and Case (i.) of Section (10) had been worked out in detail. It is the rise of correlation with more stringent limitation of opposing influences, which suggests that after all that law is true for fertility as for other characters.

by a variety of other factors of evolution. These factors are so active that the influence is reduced to ·12 of a child per generation if we take Case (i.) of Section (9), and, we have little doubt, would be practically insensible did we take all marriages without any limitation whatever. Reproductive selection must, therefore, be looked upon as always tending to increase the fertility of a race; races are not only ever tending to increase, but tending to increase the rate at which they increase—a feature not recognised by MALTHUS, but which strengthens certain of his arguments. So soon, therefore, as environment, or other circumstance, relieves the pressure of opposing factors, a race will not only increase in numbers, but also in fertility. It is this inherited character of fertility, and its constant tendency to change unless held in check by natural selection or other factor of evolution, which seems to us the source of the immense diversity in fertility to be observed not only in different species, but in local races of the same species.

III. *On the Inheritance of Fecundity in Thoroughbred Racehorses.* By KARL PEARSON, F.R.S., with the assistance of LESLIE BRAMLEY-MOORE.*

(13.) The data provided for the fertility of thoroughbred racehorses by the *stud-books*, are of a kind which cannot be hoped for except in the cases of pedigree animals kept for breeding purposes, and of specially-arranged experiments on insects, &c. We have a practically complete record of the stud-life of every brood-mare. The sire by whom she has been covered in each year is stated, and the result, barren, dead foal, living filly or colt, twins, &c., can be ascertained. It is also possible to find out whether the foal dies young, say as a yearling. By examining the whole series of stud-books the complete pedigree of any mare or sire can nearly always be found, and the correlation theoretically worked out for almost any degree of relationship.

In starting an investigation of this kind on such a great mass of raw material, it is necessary to draw up certain rules for the extraction and arrangement of data. These rules must be prepared without any definite knowledge of the character of the material in *bulk*, for this can only be found after, perhaps, some 1000 cases have been extracted and worked out. Hence the rules originally adopted are often not such as an investigator would have arranged had he known beforehand the general character of the conclusions he would reach. But the statistician cannot, like the experimental physicist, modify without immense labour his methods and repeat his experiment. The collection of his data has frequently been far too laborious a task for repetition. His raw material has been prepared in a certain manner; he may

---

* During the three years in which this investigation has been in progress, a considerable number of friends have given me substantial aid in the arithmetical work, or in the preparation of the 6,000 pedigree cards on which the results are based. Mr. BRAMLEY-MOORE has latterly been my chief helper, but I am also much indebted to Miss ALICE LEE and Mr. G. U. YULE. Miss MARGARET SHAEN and Miss LINA ECKENSTEIN have also contributed to the labour of extracting the raw data from the stud-books.

sort and rearrange his data cards in a variety of ways, but to prepare new cards on a different system is practically beyond his powers.

These remarks are made in order to meet criticism of the method in which my data cards were prepared. I could *now* possibly extract more convenient data, but that is only because of the knowledge gained in the process of examining the fecundity of several thousand horses. I did not even know, *ab initio*, the extent of variability in equine fertility; I did not even know the immense preponderance which would have to be given to certain sires, at any rate I had no numerical estimate of it. Nor had I any percentage of the number of cases in which a pedigree might end abruptly with an alternative sire.*

I saw at once that the apparent fertility of racehorses was even less close to their potential fertility (which I presume to be the inherited character) than in the case of man. Mares go at different ages to the stud, they remain—for reasons not stated—uncovered for occasional years, or periods of years; they return to the training stable for a time; they are sold abroad; they are converted into hunters, put into harness, or, as is occasionally recorded, sold to cab proprietors. This by no means invariably denotes that their fertility is exhausted; their offspring may be bad racers, or their stock unfashionable. Very frequently also we find the mare put to a cart-horse stallion for a year, a few years, or for the remainder of her career, and then no record at all is given of the result. Thus the total fertility recorded can have but small correlation with the potential fertility, and I was compelled to deal with fecundity. The insufficiency of the *apparent* fertilities, as recorded in my mare index, to solve the problem, may be illustrated in the following manner: 1100 cases of the apparent fertilities of mares and dams having had at least four coverings were tabulated (Table IX.). The following results were calculated from this table, the subscript m referring to mare and d to dam:—

$$M_d = 7\cdot6655, \qquad M_m = 6\cdot1391,$$
$$\sigma_d = 3\cdot3652, \qquad \sigma_m = 3\cdot1617,$$
$$r_{md} = -\cdot0868.$$

The probable error of $r_{md} = \cdot0202$, and thus we might argue that a fertile dam has, on the average, infertile offspring. But an examination of the above numbers shows us that the dams are more variable than the mares,† and yet the dams have been theoretically subjected to the greater selection, for they must all be granddams, or the fertility of the mares could not have been recorded. We are forced to conclude that the mares have been in some manner selected, and the form of the selection is fairly obvious on examining the table. There appears a great defect of

---

[*] Even the pedigree of such a famous racehorse as Gladiateur is soon checked by the occurrence of alternative sires. His sire, Monarque, was the son of either The Baron, or Sting, or the Emperor.

[†] The variability of mares, as a whole, not separated into mares and dams, is (see Art. 16) 3·2775.

mares in the third quadrant,[*] *i.e.*, of mares and dams of large fertility, the frequency is cut off abruptly in this quadrant. The reason for this is fairly clear. We have dealt with a limited number of years, about 30, of horse-breeding; hence, when the dam has a long record, her later offspring at any rate cannot possibly have a long one; when she has a short one, it is possible for them to have a long one. Accordingly, there has been a process of unconscious selection, which has led to a negative correlation of the apparent fertilities.

To illustrate the point further, two more correlation coefficients were obtained. In Table X. are given the apparent fertilities of mares and their dams with a minimum of eight coverings. We find:

$$M_d = 8{\cdot}6191, \qquad M_m = 7{\cdot}6309,$$
$$\sigma_d = 3{\cdot}1656, \qquad \sigma_m = 2{\cdot}8149,$$
$$r_{md} = -\,{\cdot}0876.$$

The probable error is again about ·0202. While the mares now form a group with their mean fertility almost equal to that of the dams in the previous result, their variability is markedly less. Relatively to the dams its reduction is even greater. The correlation is sensibly the same. It would thus seem that the anomalous selection of mares which thus reduces their variability so markedly below that of the dams is not in the *low* fertilities.

I now removed from the Table IX. all parts of it concerning mares with a fertility greater than 8; 867 mares and dams remained with a minimum limit of four coverings, the mares not having a greater fertility than 8 offspring. I found:

$$M_d = 7{\cdot}7636, \qquad M_m = 4{\cdot}8558,$$
$$\sigma_d = 3{\cdot}3983, \qquad \sigma_m = 1{\cdot}9887,$$
$$r = -\,{\cdot}0190.$$

The probable error of $r = {\cdot}0229$. Now the line of regression for dams on mares ought to be the same, whether we obtain it from this result or from the first results in which mares with more than 8 offspring are included. Yet, in this case, there is no sensible correlation at all. In other words, if we exclude the data for large fertilities, we should have to conclude that there was no correlation between the apparent fertilities recorded for mares and their dams. We are thus forced to conclude that apparent fertility is a character depending on the manner in which the record is formed, and must be useless for the investigation of inheritance. This investigation strengthens my *à priori* reasons for selecting fecundity, not apparent fertility, as the character to be investigated. I took the fecundity of a brood-mare to be the number of her living offspring divided by the potential number of her offspring under the given circumstances. Of both numerator and denominator of this ratio I must say a few words.

---

[*] The portion of the table cut off by vertical and horizontal lines through the means of dams and mares.

In considering the inheritance of fertility I had two different problems in my mind: (i.) Is fertility pure and simple inherited? *i.e.*, Does a very fertile mare have offspring more fertile than the average? And (ii.) What effect does reproductive selection actually have on the population? *i.e.*, To what extent is it screened by other factors of evolution; does the very fertile mare actually have more offspring than the less fertile? Is, for example, her stock weedy and likely to die early? In the case of mankind, the fertility of a woman is, as a rule, effectively brought to its limit with the end of her marriage, and accordingly I started with completed marriages. In the case of a brood-mare her effective fertility depends not on the offspring she has but on the number of these which survive foaldom. It would doubtless have been better to have treated these two problems of fertility separately, but being fairly confident from Proposition I., p. 260, that fertility must be inherited, I was more interested to test the actual effect of reproductive selection. Accordingly I selected as the numerator of my fecundity ratio, not the number of foals born, but those who survived to the yearling sales. The difference is not very great, but quite sensible. For example, the mean fecundity of 3909 brood-mares, measured in my way, = ·6343, *i.e.*, 63 surviving offspring on the average of 100 coverings.

The following table gives the result of reckoning merely barren mares and those slipping foals or giving birth to dead foals in a twenty-year period:—

AVERAGE Fecundity of Brood-mares.

Year.	Average fecundity.	Year.	Average fecundity.
1873	·712	1883	·693
1874	·703	1884	·678
1875	·707	1885	·702
1876	·697	1886	·700
1877	·692	1887	·682
1878	·680	1888	·695
1879	·683	1889	·685
1880	·666	1890	·686
1881	·680	1891	·679
1882	·667	1892	·675

The averages of five-year periods are:

·702, ·675, ·691, ·684,

and of the whole period, ·688.

There does not appear to be sufficient evidence for any secular change here, and we may take ·688 to represent the average fecundity of the brood-mare, reckoning viable offspring to the number of coverings. The difference of ·688 and ·634 gives a death-rate of 5·4 foals in 68·8, or a death-rate of 7·85 per cent. of foals before maturity. If a considerable part of this death-rate be differential, we have room for natural selection influencing the drift of reproductive selection. The standard deviation in the fecundity is, however, about ·191, or about 19 foals in the 63, or about 30 per cent.—a very great variation, so that if fecundity be inherited, a differential death-rate of the immature will hardly suffice to check it.

So much then of the numerator of my ratio. I have spoken immediately above of the denominator as if it were the number of times the mare had been covered. It is generally this, but in the relatively few cases where the mare has given birth to twins, I have counted that covering *twice*. Had this not been done the fecundity might have been greater than unity, for example even in some exceptional cases have risen to two. On the other hand, a loss of twins would have been marked by no greater change in fecundity than a loss of one foal, or the survival of *one* twin would not have been different in its effect on fecundity to the birth of a foal. In order, therefore, to avoid these difficulties—especially that of isolated individuals lying far beyond the fecundity range of 0 to 1—when twins were born the potentiality of the covering was reckoned in the denominator as two. The relative infrequency of twins causes, however, this modification of the denominator to have small influence on the result.

My next step was to form some estimate of the extent to which fecundity thus measured was the same for different periods in a mare's breeding career. I expected fecundity to diminish with age as in the case of mankind, but taking out a fairly large test number of mares, I found that their fecundity for the periods covered by two successive stud-books was in the majority of cases closely the same. With larger experience I should now lay more weight on the decrease of fecundity with age; and I also think fecundity is smaller when the mare first goes to the stud. But even thus much of the reduced fecundity of old mares seems to arise from breeders sending famous mares to the sire long after their breeding days are passed. I have several records of old mares being covered seven or eight times without offspring. This custom of breeders was much more rife in the early days of breeding than it appears to be now, when some breeders discard or sell a fairly old mare, even if she is barren two or three successive years. Clearly the custom gives the mare a fictitious fecundity, far below her real value, and probably accounts for granddams having a somewhat less fecundity than their granddaughters.

The next problem to be answered was the effect the method of forming my fecundity ratios might have on the relative numbers which would be found in different element-groups. For example, supposing the element of fecundity to be 1/10, or the element-groups 0–1/20, 1/20–3/20, 3/20–5/20, . . . 17/20–19/20, 19/20–1, would the fact that the fecundity ratio is a ratio of *whole* numbers cause, *à priori*, a greater probability of frequency in one of these element-groups than another?

To begin with, all estimation of fecundity based on less than *four* coverings was discarded. Three coverings give too rough an appreciation of a mare's fecundity, it can only fall into one of the values 0, 1/3, 2/3, and 1. The question then arises, if all the fecundities:

0/4, 1/4, 2/4, 3/4, 4/4,
0/5, 1/5, 2/5, 3/5, 4/5, 5/5,
0/6, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6,
. . . .
0/26, 1/26, 2/26, ... 25/26, 26/26,

were equally likely, how would the frequency depend on the grouping ?*

Taking 26 coverings as the probable maximum—it actually occurs—we have for the total number of fecundities given above : $5 + 6 + 7 + \ldots + 27 = 368$ separate fecundities. Let us see how they would be divided in one or two cases.

Case (i.) Let the elements be based on 1/8, or be 0–1/16, 1/16–3/16, 3/16–5/16, 5/16–7/16, 7/16–9/16, 9/16–11/16, 11/16–13/16, 13/16–15/16, 15/16–1.

The half-groups at the ends are taken so that zero and perfect fecundity should really be plotted at the middle of a 1/8 element. We find, adding up the numbers of the above fecundities which fall into the nine groups, the following frequencies :—

33·5, 42, 43·5, 44, 42, 44, 43·5, 42, 33·5.

There is thus a somewhat deficient frequency in the terminal groups, and this would probably to some extent bias the distribution.

Case (ii.) Let the elements be based on 1/15, or be

0–1/30, 1/30–3/30, 3/30–5/30, . . . 25/30–27/30, 27/30–29/30, 29/30–1.

We have the following distribution :

23, 22, 23, 23·5, 22·5, 23·5, 23·5, 22, 22, 23·5, 23·5, 22·5, 23·5, 23, 22, 23.

The bias here is only slight and the distribution is on the whole very satisfactory.

Case (iii.) Let the elements be based on 1/20, or be

0–1/40, 1/40–3/40, 3/40–5/40, . . . 35/40–37/40, 37/40–39/40, 39/40–1.

We find for the groups :

23, 13, 17·5, 17·5, 17, 18, 17, 18, 17·5, 17, 18, 17, 17·5, 18, 17, 18, 17, 17·5, 17·5, 13, 23.

Here the terminal groups have too great a frequency, and the adjacent groups too little. It is clear that the division into 1/15 elements is better than those of 1/8 of 1/20, so far as these results go. But unfortunately the different coverings do not occur in anything like the same proportions. Their exact frequencies could only be found *à posteriori,* and I was desirous of having some idea of grouping before start-

---

* Such problems are really not infrequent in statistical investigations, and seem to be of some interest for the theory of fractional numbers. Mr. FILON worked out for me the details of the cases given below.

ing the labour of extraction. I therefore weighted the different coverings on the basis of a small preliminary investigation as follows:

Case (iv.) Number of coverings,	4 to 5	inclusive,	loaded	with	2.
,, ,,	6 to 9	,,	,,		3.
,, ,,	10 to 15	,,	,,		4.
,, ,,	16 to 18	,,	,,		2.
,, ,,	19 to 26	,,	,,		1.

The resulting system of frequencies was:

54, 42·5, 45·5, 47, 45·5, 46·5, 46, 45, 45, 46, 46·5, 45·5, 47, 45·5, 42·5, 54.

This system is not so uniform as in Case (ii.). I had hoped that the 744 frequencies would have been fairly closely the double system of Case (ii.). The main irregularity occurs at the terminal groups, or those having fecundities nearly zero and nearly perfect. These I considered would be relatively infrequent, when we started with as many as four coverings, and had an average failure of about 37 in 100. The sequel showed that the assumption was legitimate, so far as regards zero fecundity, but that perfect fecundity was sufficiently frequent to cause a hump in the frequency curve for fecundity, corresponding to the group-element 29/30 to 1. The frequency of this group is greater than that of the group 27/30 to 29/30, when we start from at least four coverings. This hump entirely disappears, however, if we start with at least eight coverings. Thus I take the hump to be purely "spurious," *i.e.*, a result of the arithmetical processes employed, and not an organic character in fecundity. It depends upon our definition of fecundity, which is not a truly continuous quantity.

As the theory of correlation applied is not in any way dependent on the form of the correlation surface, beyond the assumption of nearly linear regression, the hump cannot, I think, sensibly affect our conclusions. Had I known, however, *à priori*, what the frequency of different coverings and the nature of the fecundity frequency curve would be, I should have attempted to choose such a group-element, that, with proper weighting of the coverings, there would have been no *arithmetical* bias to the terminal groups. As it was, it seemed to me that the group-element of 1/15 gave fairly little arithmetical bias—at any rate where the bulk of the frequency would occur—and it was accordingly adopted as a basis for classifying fecundities.

The dfficulty illustrates the point I have referred to, namely, that in statistical investigations the best classification can only be found *à posteriori*, but the classification adopted has usually to be selected *à priori*.

The 1/15 element being selected, the letters a, b, c, d, e, f, g, h, i, j, k, l, m, n, p, q were given to the 16 groups of fecundities from 0 to 1, as cited under Case (ii.).*

---

* A table was formed of the 368 actually-occurring fecundities, from which it was possible to at once read off the group (or it might be two groups, *e.g.*, ·5 falls half into h and half into i) into which they each fell.

Thus the fecundity of a mare was described by one of these 16 letters. Here the centre of the *j* group, for example, is ·6, and it covers all fecundities from ·56 to ·63. Thus midway between *j* and *k* we are at about the mean fecundity.

The more recent Stud-Books, vols. 12 to 17, were taken as containing more complete details and, what is more important, less in-and-in breeding, although as we shall see, this is still an important factor. These volumes cover 30 and more years of English* stud life. From these 30 years' records upwards of 5000 mares, who had been covered upwards of four times, had their fecundity ascertained. The process was a very laborious one, as each mare had generally to be sought for in several volumes, and the records in each volume are not continuous, but overlap by quite arbitrary numbers of years. Further, great care had to be taken to identify each mare properly, as the same name is very frequently repeated, and the like difficulty occurs, though to a lesser extent, in the case of sires. A card was then written, giving the name of the mare and those of her sire, her dam, and her dam's sire. Upon this card the letter indicating her fecundity was placed. A card alphabet of mares was thus formed, consisting, in the first place, of about 3000 entries. This alphabet was again gone through and the fecundity of the dams of the mares inserted on the cards till there were about 2500 cases known of mare and dam. The dams were partly found from the existing series, but it was also largely necessary to work out fresh cases. Lastly, the cards were gone through and the fecundity of the grand-dams entered in upwards of 1000 cases. This forms the first series of cards.

In the next place a card index was formed of all the sires serving during these 30 years. This contained upwards of 1000 cards. On these cards the sire's sire was entered, and the fecundity of all the mares contained in the first or mare alphabet was now taken off and placed on the card of the mare's sire. Thus the card of each sire had the letters *a*, *b*, *c*, *d*, *e*, &c., upon it, and a frequency distribution was formed on the card of each sire for the fecundity of his daughters.

The same thing was done for the sires' sires; only here recourse had again to be had to the stud-books to obtain the fecundity of the daughters of the more ancient sires. Finally, a sire-alphabet was obtained which gave the average fertility of the daughters of a sire and of the daughters of his sire, or his half-sisters. On these cards was also placed the number of mares upon which each average was based.

These two card-alphabets, the mare and sire alphabets, form the "dressed" material upon which all the subsequent calculations were based.

(14.) At this point it seems desirable to insist somewhat on the many causes which tend to make the fecundity of mares, as thus determined, to a considerable extent fictitious. Many of these were only apparent to me as I became more and more familiar with the material.

---

* Irish mares were excluded except where, for pedigree purposes, it was necessary to deal with them. Many Irish mares were further included when it came to the valuation of the fertility of mares due to a given sire.

(*a*.) Mares appear to be less fecund at the beginning and end of their breeding career. Hence, when the fecundity is based on a part only of their career, as it often must be, we do not really get a fair appreciation.

(*b*.) A more fertile mare is likely to have more daughters go to the stud than a less fertile one, and hence we get a better appreciation of the fertility of the offspring of the former than of the latter.

(*c*.) Fashion among breeders interferes largely with the exhibition of the natural fecundity of a mare. She may be a famous mare and is sent to a famous sire, even though produce is not so likely as if she were put to a sire of a different class. This appears to be practically recognised when apparently barren mares are sent in one season to two, or even three sires, or again to half-bred horses or cart-horses.

(*d*.) Brood-mares which have produced performers are kept much longer at the stud, and we have the fecundity lowered by coverings after the mare is sensibly sterile. Less important mares are removed sooner from the stud.

(*e*.) Good racing mares are often put late to the stud.

(*f*.) In a certain number of cases we are simply told that the mare had no produce for a period of years, but whether she was covered or not is unrecorded.

(*g*.) Second-rate mares, or mares thought to be near the end of their fecundity, are often sold abroad. In the latter case the fecundity is fictitiously increased; in the former we have only a short period to base it on.

(*h*.) There is no record kept of the half-bred foals, which for our purpose are as important as the thoroughbred foals. "Put to a hunter" is a not uncommon record, with no statement of the result.

(*i*.) Comparatively infertile mares, unless of very valuable stock or famous as racers, are not kept long enough at the stud to get a reliable measure of their fecundity.

(*j*.) The smaller breeders will often put mares to inferior sires, already nearly worn out, either because they own them, or because their fee is low; and thus again a full chance is not given to the fecundity of the mare to exhibit itself.

(*k*.) We have excluded in our determination of the fecundity foals dying young. This is often due to the fault of the mare, but is often again due to the environment.

(*l*.) Lastly, thoroughbred mares are highly artificial creatures, and many must suffer from their environment,* either in the matter of barrenness or slipping foal, in a manner from which the wild horse or a more robust domesticated animal would be entirely free.

These considerations may suffice to show that our values of the fecundity will only roughly represent what may be termed the natural fecundity, and we ought not for

---

* I am told that there are like difficulties with cows. Cows are very liable to slip their calves, and one cow doing so, several others in the herd will or may follow her example. There is a strong folk-belief in Wiltshire—I give it merely as evidence of what a slight change in the environment is supposed to achieve—that the habitual presence of a donkey with the herd in some way soothes the cows, and renders them less ready to slip their calves.

a moment to expect inheritance in the full intensity of the Galtonian law to be exhibited by such material.

(15.) But there is another point of very considerable importance for the weakening of correlation, namely the effect of in-and-in breeding. To get correlation we must have a diversity of parents producing a diversity of offspring, but when the parents become more and more identical, we get larger and larger arrays between which and the parents the correlation is weakened. For example: suppose the correlation found between *all* parents and offspring in the general population, and now select only all the brothers in a large array and find the correlation between them and their offspring, we shall find that the correlation is lower than in the previous case.* It would be impossible to apply theory to the present case, however, because we can only roughly appreciate the extent of such in-and-in breeding. That it is great the following statistics will show.

Of the more than 1000 sires in my sire alphabet, only 760 were sires of mares which had been covered at least four times. These 760 sires had upwards of 5000 offspring, of whom I had the fecundity recorded, but when mares with alternative sires were excluded, there remained only 4677 available mares.† These mares were distributed as follows:—

Daughters	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Above 20
Sires	280	113	78	43	29	22	20	21	22	14	10	10	11	11	8	6	2	6	4	4	46

Here the second line gives the number of sires having the number of daughters in the first line in the 4677 cases, which I take to be a fair sample.

Thus over a third of the sires had only one mare. Two-thirds of the sires had together only one-fifth of the mares. *Seventy-six* of the sires were fathers of about half the mares, and 46 sires alone produced 1801 mares, almost as many as 642 sires did. We are here dealing with the fairly long period of 30 years, but even making due allowance for young stallions commencing and old stallions concluding their stud career, it will be manifest that our sample shows that the great bulk of mares for the period in question were the offspring of comparatively few sires.

But let us look at the problem from the standpoint of the sires. My 760 sires

---

* The theory of such cases is fully developed in a memoir on the influence of selection on correlation not yet published.

† Some other cases were also excluded for diverse reasons.

were *all* fathered among themselves except in 49 cases. In other words, they were the sons or grandsons of only 49 sires. Of these 49 sires, there were 12 whose pedigree I could not trace,* but they were very probably sons of sires already on my list or among the remaining 37. In the majority of cases they appeared only as the sire of one stallion. The remaining 37, whose pedigree I could trace, were descended at once or in very few generations from 9 sires.† Thus both from the standpoint of the mares and of the sires we are dealing with a closely in-bred stock, and this is one and probably a very important factor in the weakening of the fecundity correlation.

Having regard to these difficulties, if we can succeed in showing that fecundity in thoroughbred racehorses is inherited, we can be fairly confident that we have only reached a lower limit of the correlation coefficient.

(16.) *On the Inheritance of Fecundity in the Female Line.*

(i.) A preliminary investigation must here be made, in order to determine the ρ of the formulæ given in Proposition III. (p. 269) we want the correlation of fecundity with fertility. If ϕ be the fecundity, f the apparent fertility, and c the number of coverings, twins counting as a double covering, we have:

$$\phi = f/c,$$

whence if we determine the correlation between ϕ and f, numerous constants will follow. Table XI. gives the correlation between fertility and fecundity for 1000 brood-mares. We found:

$$M_\phi = \cdot 6375, \qquad M_f = 6\cdot 515,$$
$$\sigma_\phi = \cdot 1810, \qquad \sigma_f = 3\cdot 2775,$$
$$\rho = r_{\phi f} = \cdot 5152,$$
$$v_\phi = 100\sigma_\phi/M_\phi = 28\cdot 39, \qquad v_f = 100\sigma_f/M_f = 50\cdot 31,$$

where v_ϕ and v_f are the "coefficients of variation."‡ Here by YULE's Theorem § $r_{\phi f}\sigma_\phi/\sigma_f$ is the slope of the line which most closely fits the curve of regression for fecundity on fertility. If we supposed this curve to be straight, then the line must coincide with it. Now since fecundity vanishes with fertility, the curve passes through the origin, and hence, if the regression be linear, the line must also pass through the origin. In this case, as is shown on p. 270, $r_{\phi f} = v_\phi/v = \cdot 5644$. The difference between ·5644 and ·5152 may be taken, as it is several times the probable error, to indicate that the regression curve between fecundity and fertility is only approximately linear.

The variations in both fertility and fecundity are here large. Accordingly we

---

* Stockmar, Sovereign, Andover, Phaeton, Prince Caradoc, Robert Houdin, Pylades, King of Kent, Garry Owen, Calaban, Homily and Taurus.

† Tramp, Sir Peter Teazle, Catton, Buzzard, Orville, Diomed, Sorcerer, Dr. Syntax, Marske.

‡ 'Phil. Trans.,' A, vol. 187, p. 276.

§ 'Roy. Soc. Proc.,' vol. 60, p. 477.

must use the formula (i.) for the mean value of an index given in my memoir on spurious correlation.* We shall then obtain an approximate value to the mean number of coverings of each mare. Formulæ (iii.) of the same paper will then give the standard deviation for the number of coverings. In our present notation:

$$c = f/\phi,$$

and therefore:

$$M_c = \frac{M_f}{M_\phi}\left(1 + \left(\frac{v_\phi}{100}\right)^2 - r_{f\phi}\left(\frac{v_f v_\phi}{10,000}\right)\right),$$

$$\sigma_c = M_c \sqrt{\left(\frac{v_f}{100}\right)^2 + \left(\frac{v_\phi}{100}\right)^2 - 2r_{f\phi}\frac{v_f v_\phi}{10,000}}.$$

We find:

$$M_f/M_\phi = 10{\cdot}2196,$$

and:

$$M_c = 10{\cdot}2196 \times 1{\cdot}007 = 10{\cdot}2911,$$

$$\sigma_c = 4{\cdot}4455,$$

$$v_c = 43{\cdot}20.$$

To the same degree of approximation we can further ascertain the correlations between the number of coverings and the apparent fertility and fecundity, *i.e.*, r_{cf}, and $r_{c\phi}$. A short investigation similar to those in the memoir on spurious correlation just cited shows us that:

$$r_{cf} = (v_f - r_{f\phi}v_\phi)/v_c,$$

$$r_{c\phi} = (r_{f\phi}v_f - v_\phi)/v_c.$$

These lead to the numerical results:

$$r_{cf} = {\cdot}8259, \qquad r_{c\phi} = -{\cdot}0572.$$

The conclusions to be drawn from these results are all of some interest. In the first place we may ask: How does M_ϕ agree with its value found from other and more complete series? For 4677 mares—my complete series without mares with alternative sires—the average fecundity was ·6373. A better agreement could not have been hoped for. In a group of 1509 mares dealt with for variation only and entered as "daughters" on the cards—so that they had not been selected by the fact that their daughters must have recorded offspring, as is the case with "dam" entries—I found the following results:—

VARIATION in Fecundity of 1509 Brood-mares (Four Coverings).

Fecundity.	*a.*	*b.*	*c.*	*d.*	*e.*	*f.*	*g.*	*h.*	*i.*	*j.*	*k.*	*l.*	*m.*	*n.*	*p.*	*q.*
Frequency	9	3	11	26	46	43·5	85	122·5	154·5	232·5	194	223	146	100	23	90
Total 1509. $M_\phi = {\cdot}6345$. $\sigma_\phi = {\cdot}1965$.																

---

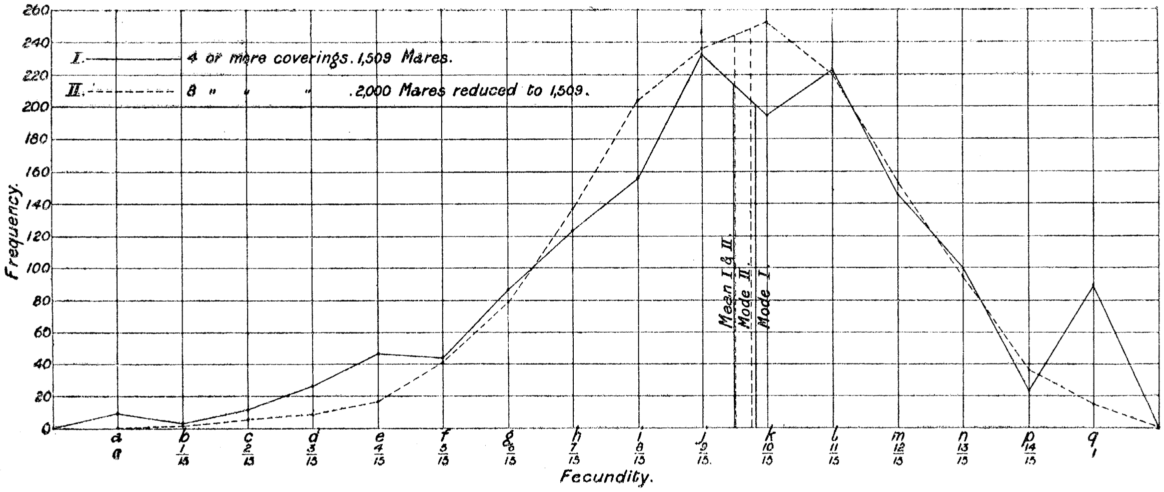
* 'Roy. Soc. Proc.,' vol. 60, p. 492.

Now this is precisely what we might expect; the mares belonged to a class, of which we are not certain whether their daughters have or have not recorded fecundity. The mean fecundity is therefore decreased and the variability increased. Add to this group 2400 mares, all of which had had their daughters' fertility recorded, and we find for 3909 mares, $M_\phi = \cdot 6345$ and $\sigma_\phi = \cdot 1910$, *i.e.*, the mean fecundity ascends and the variability falls. Illustration of this law will be found in the following two groups:—

	M_ϕ.	σ_ϕ.
1200 mares	·6337	·1888
1200 dams	·6525	·1643

Thus we send up the mean fertility and lower the variability by separating into two groups the pedigree of one which has a longer record. This is precisely in accordance with the theory already developed. Our mean fecundity and variability for brood-mares may be considered as constant characters, and variations in their values beyond their probable errors due to conscious or unconscious selection in the record itself, or in our extracting from it.

The reader will notice at once, if he turns to the diagram of the above frequency, (i.) that there is a small hump at (a) of no practical importance, and a larger one at (q), perfect fertility being fairly frequent with only four coverings, and there being from the arithmetical processes involved a bias towards (q) as compared with (p). (ii.) The distribution of frequency, although somewhat ragged, is quite clearly not normal, but of the character which in other papers I have called *skew*. Were there any occasion, it would be easy to fit it with one of my skew curves. To mark how (i.) will disappear and (ii.) become still more apparent, I have placed on the diagram the frequency distribution for 2000 mares reduced to the same scale.



VARIATION in Fecundity of 2000 Brood-mares (Eight Coverings).

Fecundity	*a.*	*b.*	*c.*	*d.*	*e.*	*f.*	*g.*	*h.*	*i.*	*j.*	*k.*	*l.*	*m.*	*n.*	*p.*	*q.*
Frequency	0	2	7·5	11·5	21·5	55	104·5	182	271·5	315	337	293·5	204	127	49	19
Total, 2000. $M_\phi = \cdot 6330$. $\sigma_\phi = \cdot 1568$.																

Thus, making the minimum number of coverings 8 instead of 4, has removed the terminal humps, zero fecundity is now unknown, and perfect fecundity very rare. We have reached a smooth skew frequency distribution ; we see fecundity as a continuous character obeying the usual laws of variation.* The mean fecundity in the two cases is sensibly the same, ·633, but owing to the fact that we have made a selection of a limited group in the second case, the variability is considerably decreased.

The average apparent fertility of brood-mares, 6·515, must not be confused with their average real fertility, for, as we have seen, we have in many cases not a complete record of their stud-life, or such a full record has not been used (*e.g.*, in case of mares still at the stud, but having been already covered four or more times). Its 50 per cent. variation shows that an apparent fertility of 9 to 12 is not infrequent. The average number of coverings being 10 and more, it will be seen that the records of between 50,000 and 60,000 coverings have been dealt with to form our mare and sire alphabets. The large variability in the number of coverings shows that 15 to 20 coverings will not be infrequent, and cases of 26 actually occurred. Lastly, we have the correlation between fertility and the number of coverings, high as might be supposed, for a high apparent fertility could only be exhibited by many coverings. Although a low apparent fertility might correspond to any number of coverings, still, in practice a sterile mare will not be sent indefinitely to the sire. The correlation between the number of coverings and the fecundity is small and *negative* (– ·0572). This follows from the principle that, fertility being the same, a high number of coverings reduces the fecundity, and this factor is more potent than the high correlation of fertility and the number of coverings.

(ii.) Table XII. exhibits the correlation of 1200 mares and their dams with regard to fecundity. Here the more fertile dams are weighted with their fertility, and at least four coverings were required of each mare. If the subscript m refers to mare, and d to dam, we find :

---

* The actual equation to the curve referred to the mode ·6531 as origin, the axis of x being positive towards perfect fecundity, and the unit of x being 1/15 is:

$$y = 342{\cdot}187\ (1 + x/47{\cdot}1358)^{82{\cdot}6261}\ (1 - x/12{\cdot}1106)^{21{\cdot}2291}.$$

The fit will be found to be very satisfactory.

$$M_m = \cdot 6337, \qquad M_d = \cdot 6525,$$
$$\sigma_m = \cdot 1888, \qquad \sigma_d = \cdot 1643,$$
$$r_{md} = \cdot 0831.$$

The coefficient of regression = ·0945.

The probable error of the correlation is ·0193 and of the regression[*] ·0195. Thus these quantities are four to five times their probable errors, and we conclude that fecundity is certainly inherited.

The intensity is far below that suggested by the law of ancestral heredity, but it nevertheless exists. Its lowness is probably due to the fictitious character of the fecundity owing to the causes indicated on pp. 298–9. An attempt must now be made to eliminate some of the factors disguising the fecundity, but to do so is by no means so easy as in the case of fertility in man.

(iii.) My first idea was that by taking a higher limit to the number of coverings a closer approach might be obtained to the true, *i.e.*, the inherited fecundity. Accordingly Table XIII. was formed for the correlation of 1000 mares and their dams, when the minimum number of coverings was eight. But I did not recognise that this would give far greater weight in the Table to the older mares, and that accordingly causes (*d*) and (*i*) of p. 298 would now play a much larger part in disguising the true fecundity than before. There appears to be no limit to the number of times a famous old mare may go to the stallion when there is very small hope of any offspring.

Table XIII. gives us the following results:

$$M_m = \cdot 6300, \qquad M_d = \cdot 6360,$$
$$\sigma_m = \cdot 1633, \qquad \sigma_d = \cdot 1500,$$
$$r_{md} = \cdot 0652.$$

The coefficient of regression = ·0708.

The probable error of the correlation is ·0212, and of the regression ·0213, both less than a third of the observed values. We should again conclude from this result that fecundity is inherited, although it offers less strong evidence than the previous case. The influence of selection[†] is at once apparent in the great reduction of the variabilities. The fact that we are throwing the determination of fecundity more on to the old age period of life appears from the reduced mean fecundities. I attribute the reduction in the fecundity-correlation to this source, *i.e.*, the very diverse treatment which old mares receive at the hands of different breeders.

(iv.) I made another attempt to remove screening causes by taking 1200 more

---

[*] Pearson and Filon: 'Phil. Trans.,' A, vol. 191, p. 214.

[†] The effect of such a selection as the above in reducing correlation is dealt with in my paper on the influence of selection on correlation.

mares, not identical with the series in (ii.)* and working out their dams' records most carefully, rejecting any cases in which the breeder was clearly sending the mare to the stallion long after it was obvious (*post facto*) that she was sterile. In this case four coverings were retained as a minimum, and the results are given in Table XIV. We find :

$$M_m = \cdot 6369, \qquad M_d = \cdot 6616,$$
$$\sigma_m = \cdot 1885, \qquad \sigma_d = \cdot 1604,$$
$$r_{md} = \cdot 0995.$$

The coefficient of regression = ·1169.

The probable error of the correlation is ·0193, and of the regression ·0194 ; the correlation is accordingly more than five, and the regression more than six times its probable error. We conclude that fecundity is most certainly inherited. The regression found is, however, only about two-fifths of what is required by the law of ancestral heredity,

(v.) It has been suggested that fertility or fecundity might alternate in *two* generations ; when the offspring are numerous their offspring might have less fertile or fecund offspring. I do not see how this would be possible without its exercising an influence on the correlation of two generations, for we must come to one fertile followed by an infertile generation. But I had made preparations in my alphabet of mares for testing the correlation between mares and their granddams, and I went on to the construction of a table, although the results for mares and their dams showed me that whatever result might be reached, it would be within the probable error of the observations. I reached this conclusion in the following manner : If we go back one generation we introduce, owing to the nature of the record, so much fictitious correlation and so much in-and-in breeding that the coefficient of inheritance is reduced to two-fifths or less of what its value should be according to the law of ancestral heredity. In going back two generations we come to fewer mares, to more in-and-in breeding, and to just the type of famous old mare, whose breeder kept her at the stud long after she was sterile. I expected accordingly a great and artificial fall in the fecundity of granddams and a double drop, something like $\frac{2}{5} \times \frac{2}{5}$, in the value of the regression as indicated by the law of ancestral heredity. This would reduce the apparent regression to about $\frac{2}{5} \times \frac{2}{5}$ of ·15, or to about ·025, say, a value about equal to the probable error of the table. The results actually reached are given in Table XV., and we find, if the subscript g refer to granddam :

$$M_m = \cdot 6345, \qquad M_g = \cdot 6232,$$
$$\sigma_m = \cdot 2040, \qquad \sigma_g = \cdot 1687,$$
$$r_{mg} = \cdot 0169.$$

The coefficient of regression = ·0204.

---

[*] In the first series the mares' names run from A to G ; in the second from G to M, with 300 additions made to the A to G series, while I was completing my alphabet.

The probable error of the correlation is ·0213, and of the regression ·0213. Thus these results are not significant in themselves, but they are exactly what we might expect on the above hypothesis. Taken with the other five tables which we have worked out for the inheritance of fecundity, they are significant, for every one of them gives a *positive* correlation, however small it be, and thus adds to the accumulated evidence that fecundity is a heritable character.

(vi.) It remains to test our results by the theory developed on pp. 269 *et. seq.* But a difficulty comes in here. Turning to (xviii.) and (xix.) on p. 268, we cannot feel justified in putting $M_1 = M_2$, for there is a secular difference in the fecundity of mares and dams, owing to the fecundity of the older brood-mares being based on a longer period and liable to the disturbing causes so markedly manifest in the correlation of mares and granddams (see my remarks, p. 305). If we combine (xviii.) and (xix.) we find

$$M''_2 - M'_2 = r\,\frac{\sigma_2}{\sigma_1}(M''_1 - M'_1).$$

Now r is small, and it will accordingly be legitimate to put $M'_1 = M'_2$ and $\sigma_1 = \sigma_2$ on the right, we have then

$$(M''_2 - rM''_1)/(1 - r) = M'_2.$$

From this we deduce for the results in (ii.) on p. 304

$$M'_2 = \cdot 6321.$$

Turning now to (xix.), it may be written

$$M''_1 = M'_1 \left\{1 + \frac{\sigma_1^2}{M_1^2}\left(\frac{\sigma'^2_1/\sigma_1^2}{1 + \sigma_1^2/(\rho^2 M_1^2)}\right)\right\}.$$

The second term in the curled brackets is small, and in it we may put to a first approximation $\sigma'_1 = \sigma_1 = \sigma_2$ and $M_1 = M'_2$. We then have

$$M''_1 = M'_1 \left\{1 + \left(\frac{\cdot 1888}{\cdot 6321}\right)^2 \left(\frac{1}{1 + \left(\frac{\cdot 1888}{\cdot 6321}\right)^2 \frac{1}{(\cdot 5132)^2}}\right)\right\},$$

or,

$$M''_1 = M'_1 \times 1\cdot 0666.$$

Substituting the value of M''_1 we find

$$M'_1 = \cdot 6118.$$

We thus see a difference in the fecundities of the unweighted dams and unweighted mares of ·6118 and ·6321, or about 2 foals more in the hundred appear to survive in the later generation. This is very probably due to the causes already indicated as affecting the apparent fecundity of the older mares (see p. 298). The influence of reproductive selection changes these quantities to ·6337 in the case of the daughters, and to the *apparent* high fecundity of ·6525 in the case of the dams.

We can now find σ_2 to a second approximation by aid of (xxi.). In the small term multiplied by r, we put $\sigma_1 = \sigma_2 = \sigma''_2$. Hence we find

$$\sigma_2^2 = \sigma''^2_2 + r^2 (\sigma''^2_2 - \sigma''^2_1),$$

and deduce, on substituting the numerical values,

$$\sigma_2 = \cdot 1896,$$

or is scarcely different from σ''_2. We accordingly conclude that we may quite reasonably assume the variability of the mares to represent the variability of the mares without reproductive selection, but the effect of weighting the dams with their fertility is to reduce the variability of the dams from about ·1896, if there be no secular change, to an apparent value as low as ·1643.

The same formulæ applied to the slightly better results in (iv.) on p. 305 give us:

$$M'_1 = \cdot 6205, \text{ and } M'_2 = \cdot 6342.$$

If we pass back from M'_1 and M'_2 to M_1 and M_2 we find:

	First case.	Second case.
M_1	·5460	·5567
M_2	·6266	·6278

If these results be considered as valid, we notice a remarkable difference between the fecundity of the younger and elder generation. While the crude results on pp. 304 and 305 might lead us on first examination to suppose the elder generation more fecund than the younger, these results show us that it is distinctly less so. The greater part of the difference, however, is due, not to a secular change, but to the causes we have so often referred to as weakening the fecundity recorded for the older mares. At the same time the whole system of breeding is so artificial that we may well doubt whether our equations (i.) and (v.) can be legitimately applied. For the chance of a mare getting into the stud-book as a dam, *i.e.*, having daughters at the stud, depends less on her fertility than on the degree of fashion in her stock. Thus the record weighting with fertility is hardly a probable hypothesis, and the values just given for M_1 are, I suspect, much below what they should be. For the above reason I have not proceeded to consider the changes in variability connoted by (ii.) and (xxii.). As I have made no attempt to form a correlation table for mares and dams in which the dam would have only one daughter to her record, I cannot make any plausible guess at the real magnitude of the cubic summation term in (xxii.). Apart, however, from the numerical application of these variation formulæ to a somewhat doubtful case, we see in these formulæ the theoretical basis for the observed fact that the fecundity of mothers is far less variable than that of daughters. It is really only an apparent divergence, due to the fact that the mothers have been weighted with their fertility; this, while it increases the apparent mean of their fecundity, reduces its apparent variability.

(17.) *On the Inheritance of Fecundity in the Brood-mare through the Male Line.*

For the thoroughbred horse this problem is fairly easily answered by investigating whether mares related to the same stallion have any correlation between their fecundities. The two cases I have selected are: (i.) "Sisters," daughters of the same sire, but in general not of the same mare; and then (ii.) "Nieces" and "Aunts," or daughters of a sire and the daughters of his sire. As we have only 760 sires and nearly 5000 mares, the daughters or aunts fall into rather large arrays, and we are compelled to use the methods discussed in Proposition IV., A and B. Even so the arithmetical work for a correlation based on the index of sires was far more laborious than for one based on the index of mares.

(i.) *To find the Correlation between Half-Sisters, Daughters of the same Sire.*

Here we have to use formulæ (xxiii.), (xxv.), and (xxvi.) of pp. 272–273. In order to do this a table was formed of the mean fecundity M of the array of sisters due to each sire, and of $\frac{1}{2}n(n-1)$, the number of pairs of sisters in each array. Then the products $\frac{1}{2}n(n-1)$ M and $\frac{1}{2}n(n-1)M^2$ were formed, and the numerator of (xxiii.), or σ_a^2, calculated by adding up for all the 760 sires. The result gave:

$$\sigma_a^2 = \cdot 6655167,$$

where the unit is the fecundity group element of 1/15. The number of pairs of sisters dealt with was 54,305. The denominator $\sigma_0^2(1-\rho^2)+\sigma_a^2$ is not so easily ascertained. σ_0 is the standard deviation of all the series of mares who are sisters without weighting; $\sigma_0\sqrt{(1-\rho^2)}$ is the standard deviation of an array of sisters, or if the regression be not linear, the mean of such standard deviations for all arrays, or rather its square is the mean of the squares of such standard deviations; ρ is the correlation between a patent character in the daughter and a purely latent character in the sire, and cannot therefore be found directly.

In order to get an appreciation of the standard deviation of an array of sisters—it being practically impossible to work out these quantities for 760 arrays—I selected twenty sires having fairly large arrays of daughters, and reached the following results:

TABLE of Arrays of Mares, which are Half-Sisters.

Sire.	No. of mares.	Mean fertility.	S.D. of array.
Speculum	76	9·697	2·989
Sterling	52	10·750	2·545
Scottish Chief	67	9·201	3·176
Newminster	64	8·875	2·497
Parmesan	37	9·708	3·076
Macaroni	81	10·210	2·770
King Tom	53	9·689	2·748
Lord Clifden	41	9·878	2·086
Hermit	79	9·437	3·003
Blair Athol	87	9·057	2·752
Lord Lyon	32	9·125	3·314
The Duke	35	9·186	2·474
Doncaster	37	9·297	3·021
Adventurer	58	10·466	2·621
Cathedral	43	9·267	2·847
Rosicrucian	59	10·932	3·094
Stockwell	80	9·131	2·093
Rataplan	40	8·222	2·201
Y. Melbourne	55	9·064	2·301
Thormanby	41	9·951	1·651
Totals	1117	191·143	53·259
Mean	55·85	9·55715	2·66295
Ditto in actual units*	. .	·6371	·1775

I next took the mean and standard deviation of the 1117 mares to obtain σ_0. The mean fecundity was now found to be 9·5685 and $\sigma_0 = 2·7824$, or in actual units ·6379 and ·1855. Clearly only about $\frac{1}{8}$ per cent. difference is made whether we take the mean fecundity of the 1117 mares, or the mean of the unweighted means of the twenty arrays. Knowing σ_0 and $\sigma_0\sqrt{1-\rho^2}$ we can now find ρ. We have almost at once

$$\rho = ·2900.$$

This is probably the first determination of a coefficient of inheritance between a latent character in one sex and a patent character in the other sex. We see that it has almost exactly the value required (·3) by the law of ancestral heredity, or we conclude, *mares inherit from their sires a fecundity governed closely by the law of ancestral heredity.*

If the reader asks why is not the intensity reduced in this case in the same manner that we find it reduced in the case of the inheritance from the dam, the reply is :

(i.) In the case of the dam and mare, *both* quantities to be correlated are liable to fictitious values. In the case of sire and mare, we deal with only one.

---

* A fecundity unit is taken to be 1/15, for this is the unit of grouping.

(ii.) The influence of fictitious values has been shown on pp. 276–277 to chiefly affect the coefficient of correlation and not the standard deviation.

Now the present result is based solely on the calculation of standard deviations, or on the variability of fecundity as a whole and in arrays. It is accordingly not influenced to nearly the same extent by the existence of fictitious values. Could we calculate the variability of the arrays of daughters due to individual mares, we should probably get a better result for inheritance in the female line.*

The above result is so satisfactory that I have little doubt that we have determined a very good value for $\sigma_0\sqrt{1-\rho^2}$. Substituting it we find for the correlation between half-sisters:

$$r = \frac{\cdot 66552}{7\cdot 09130 + \cdot 66552} = \cdot 0858.$$

The law of ancestral heredity gives for *half*-sisters $r = \cdot 2$, and $\frac{2}{5}$ of this $= \cdot 08$.

Thus we see that the collateral heredity between half-sisters, daughters of the same sire, is quite sensible, and is almost what we might have predicted would be the result, if we supposed correlation to be weakened, as in the previous cases, to $\frac{2}{5}$ of its value by fictitious records.

It is worth while to consider the amount of fictitious fecundity suggested by the reduction factor $\frac{2}{5}$. We have only to suppose the n_1/N of our p. 277 to be $\frac{2}{5}$. Now we may well assume the chance of a fictitious fecundity being recorded to be the same for either one of a pair of sisters; hence we shall have $p = q$, and therefore, from the result on p. 276, we find $(p-1)^2/p^2 = \frac{2}{5}$. This gives us $(p-1)p = \sqrt{\cdot 4}$, and $(n_1 + n_3)/\mathrm{N}$ the fraction without fictitious values $= (p-1)/p = \cdot 6325$. Thus in order to introduce the reduction factor of $\frac{2}{5}$ by the occurrence of fictitious values of the fecundity, we should have to suppose about 37 per cent. of fictitious values to occur. This is, of course, a sort of average; many values will probably be only partially fictitious, *i.e.*, will to some extent approximate to their real values. Considering the very artificial character of the thoroughbred brood-mare, and the uncertainty of her treatment by breeders, this does not seem such an immense percentage that it would force us to the conclusion that the law of ancestral heredity cannot be true for the inheritance of fecundity.

(ii.) *To find the Correlation in Fecundity between the Sisters of a Sire and his Daughters.*

What we want is really the correlation between aunts and nieces, but they

---

* The standard deviations for the arrays of mares in Table XII. were indeed worked out for the twelve cases of dams from e to q. The mean of these cases was sensibly the same whether the simple mean, or the mean weighted with the numbers in the array was taken, and equalled 2·8091 or ·1823. This is $\sigma_0(1-r^2)^{1/2}$. But by p. 48, $\sigma_0 = \cdot 1888$, whence we deduce $r = \cdot 1375$, and the regression equals ·1581. Thus we have found a substantially larger value for r than that on p. 304 by dealing with variabilities, and not direct correlations. This gives additional evidence, if any were needed, of the inheritance of fecundity.

TABLE XVI.—Correlation Table for Weighted Mean Fecundities of Ar

			c		d		e		f		g	
			3–3·5.	3·5–4.	4–4·5.	4·5–5.	5–5·5.	5·5–6.	6–6·5.	6·5–7.	7–7·5.	7·5–8.
Stallion's sire.	e	4·5–5	..	..	..	..	..	..	..	..	..	..
		5–5·5	..	..	..	..	..	..	..	..	..	..
	f	5·5–6	..	..	..	..	..	..	..	..	..	..
		6–6·5	..	..	..	..	..	..	..	..	..	..
	g	6·5–7	..	..	..	..	..	·5	·5	..	..	..
		7–7·5	..	..	..	..	..	·5	·5	..	..	..
	h	7·5–8	..	..	..	..	..	..	..	..	..	..
		8–8·5	..	..	..	..	..	..	..	4	4	44·25
	i	8·5–9	..	..	..	..	..	..	..	22	22	57·25
		9–9·5	9	9	..	20	49	37·5	8·5	8·5	145	18
	j	9·5–10	3	3	..	66	66	80·5	64·5	1·5	137	179·5
		10–10·5	..	..	..	107	107	74·5	56	127	221·5	428·5
	k	10·5–11	..	3	3	19·5	40	243	10·5	22	22	76
		11–11·5	29·5	29·5	..	6·75	35·75	29·75	·75	30	18	301
	l	11·5–12	3·5	3·5	..	·75	6·25	13·25	7·75	23·75	23·75	44
		12–12·5	..	..	..	..	..	..	..	2·25	2·25	..
	m	12·5–13	..	..	..	..	..	..	..	..	..	..
		13–13·5	..	..	..	..	..	..	..	1·5	1·5	..
	n	13·5–14	..	..	..	·25	·25	..	..	..	..	..
		14–14·5	..	..	..	·25	·25	..	..	..	..	..
	p	14·5–15	..	..	..	..	..	..	..	..	..	..
		15–15·5	..	..	..	..	..	..	..	..	..	..
	q	15·5–16	..	..	..	..	..	..	..	..	..	..
		16–16·5	..	..	..	..	..	..	..	..	..	..
Totals ..		..	45	48	3	220·5	304·5	479·5	149	242·5	597	1148·5

rrays of Stallion's Daughters and Stallion's Sire's Daughters. 138,424 cases of Aunts a

Stallion.											
		i		*j*		*k*		*l*		*m*	
–8·5.	8·5–9.	9–9·5.	9·5–10.	10–10·5.	10·5–11.	11–11·5.	11·5–12.	12–12·5.	12·5–13.	13–13·5.	13·5–14.
·25	·25	..	..	..	..	..	..	..	..	..	·25
·25	·25	..	..	..	..	1	1	..	..	..	·25
..	..	..	60·25	·25	..	1	1	..	..	..	..
·75	·75	..	·25	3·25	..	..	1·5	1·5	..	..	..
·75	·75	..	..	3	..	..	1·5	1·5	..	..	..
..	2	2	..	..	..	..	..	..	..	..	..
28	24	..	4·25	16·25	..	..	..	..	..	..	..
79·25	31	47·5	8·25	15·75	110	..	..	1·5	1·5	·25	·25
315·25	328	126·5	241·75	2162·75	1·5	396	5	30·5	163·5	18·25	·25
15	332·5	276	217·75	770·25	3603	317·5	196·5	226·5	22	31·5	9·5
235·25	874·5	1415·25	3782·5	9020	5316	5197·5	1702·5	274·5	139	459·5	4
357·75	1903	3554·25	3945·5	24,628	7165	7030·5	6263·5	640·5	152	202	161·5
224	705·5	676	2482·5	8148	2205·25	1502·75	1383	2122·5	406·75	72·25	470·5
67	245·5	747	526	944	1862·25	3692·25	3283·5	214	145·75	173·25	164
99	86·25	367·25	1476·25	1347·75	1113·75	1877·25	343·5	263	72·5	20·5	4
7·5	25·75	·75	2·75	7·25	106·75	3·75	..	1	..	..	..
..	..	..	..	..	1·5	1·5	3	3	..	32	..
7·5	1·75	1	2·5	1·75	..	..	22·75	1·75	..	..	..
..	·25	1	1	·25	..	..	27·25	·25	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	·5	·5	..	..
..	..	..	..	..	..	..	..	·5	·5	..	..
1437·5	4562	7214·5	12,751·5	47,068·5	21,485	20,021	13,235·5	3783	1104	1009·5	814·5

unts and Nieces. (See p. 311.)

	n		*p*		*q*	
13·5–14.	14–14·5.	14·5–15.	15–15·5.	15·5–16.	16–16·5.	Totals.
·25	·25	..	..	..	..	1
·25	·25	..	..	..	..	3
..	..	..	..	..	..	62·5
..	..	..	..	..	..	8
..	..	..	..	..	..	8·5
..	..	..	..	..	..	5
..	..	..	..	2	2	76·5
·25	..	..	..	2	2	351·5
·25	11	11	..	..	..	3912·5
9·5	..	..	..	20	20	6362·5
4	3	2	..	32	32	29,090·5
161·5	5·5	42	..	114	114	57,400·5
470·5	14·5	..	..	73·5	73·5	20,999·5
164	65·5	·5	·5	5·5	5·5	12,623
4	..	7·5	7·5	10	10	7232·5
..	..	..	..	1	1	162
..	..	..	..	..	..	41
..	..	..	..	3	3	48
..	·75	·75	..	..	..	32
..	·75	·75	..	..	..	2
..	..	..	..	..	..	0
..	..	..	..	..	..	0
..	..	..	..	..	..	1
..	..	..	..	..	..	1
814·5	101·5	64·5	8	263	263	138,424

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

are not "aunts" and "nieces" in the human sense, for the aunts are only half-sisters of the sire. By a process similar to that on pp. 408 and 409 of my paper on the "Law of Ancestral Heredity,"* I deduce that the correlation between a sire's sisters and daughters ought to be ·05, and not ·15 as in the case of Man. If this be weakened down to the $\frac{2}{5}$ of previous results, we should not expect a result differing much from ·02. As the variability of the elder generation is always less than that of the younger, we ought to expect a coefficient of regression of about this value. The theory used will be that of p. 273 of the theoretical part of this paper. The weighted mean fecundity found for the arrays of aunts and nieces was as follows :—

	Without grouping.	With grouping.
Arrays of aunts	·6195	·6199
Arrays of nieces	·6346	·6338

The grouping was done in fecundity units of $\frac{1}{2}$, *i.e.*, 1/30 change in fecundity. The agreement may accordingly be considered very good. The "aunts" are the daughters of the older sires, who owing to in-and-in breeding form a comparatively small group, and are the sires of mares belonging to the older period, whose fecundity is much weakened by causes already referred to. Their mean fecundity is slightly less than that of granddams, given on p. 305, while the mean fecundity of their nieces agrees well with that for the corresponding group of mares.

The method of grouping being adopted, a correlation table was formed for the mean fecundities of arrays of mares, daughters of a sire, and of arrays of mares, daughters of his sire. This is Table XVI. Here each mean is weighted with the number of pairs of aunts and nieces in the two arrays, *i.e.*, the extent of the data on which it is based. It represents accordingly 138,424 pairs of aunts and nieces. The following results were obtained, corresponding to 687 pairs of sires :—

Sire's Sire.	Sire.
$M_a = \cdot 6199.$	$M_{a'} = \cdot 6338.$
$\bar{\sigma}_a = \cdot 04344.$	$\bar{\sigma}_{a'} = \cdot 07609.$

$$R = \cdot 1174.$$

It will be at once noticed how much more variable are the array-means for the sire than for the sire's sire. The means of many of the sire's arrays are based upon small numbers, which would have been selected out, if we had gone to another generation as in the case of the sire's sire.

It will clearly not be legitimate in this case to put $\bar{\sigma}'_a = \bar{\sigma}_a$ as suggested on p. 274. There is probably no secular change of importance here, but the sire's sire requiring

---

[*] 'Roy. Soc. Proc.,' vol. 62.

three generations from the record is really more stringently selected than the sire with only two. We can now form $\bar{\sigma}$ and $\bar{\sigma}'$ by (xxviii.) and (xxix.), if we adopt suitable values of $\bar{\sigma}_0$ and $\bar{\sigma}'_0$, ρ, as we have seen, may with high probability be put equal to ·3 (p. 309). σ_0 for groups of daughters, on p. 309, is given as ·1855, but since this certainly included a fair number of what are now aunts, it must be somewhat too low for $\bar{\sigma}'_0$. We can well put $\bar{\sigma}'_0$ equal to the ·1888 of the mares on p. 304. $\bar{\sigma}_0$ for aunts cannot be as low as the standard-deviation of dams on that page, as many of the aunts may never appear in the record as granddams,* *i.e.*, they are less stringently selected. The mean of the two results for mares and dams may, perhaps, be taken as a close enough approximation for our present purpose, or $\bar{\sigma}_0 = ·1765$. We then deduce

$$\bar{\sigma} = ·1739, \qquad \bar{\sigma}' = ·1955.$$

If we compare the results now found with those for sisters cited on pp. 308 and 309, we find :—

	"Aunts."	"Sisters."	"Nieces."
M	·6199	·6371	·6338
σ_a	·0434	·0544	·0761
σ'	·1739	·1855	·1955

The accordances and divergences are much what we might expect, except in the case of σ_a. We should, *à priori*, have expected "sisters" to have approached nieces more nearly than aunts. The work has been gone carefully through, but I have not succeeded in finding any error. In the "nieces," of course, the weighting of an outlying fecundity-mean due to a sire with but few daughters, may still be large, if his sire have numerous daughters; this cannot occur in the case of "sisters," as the weighting depends only on the number in the array. The like heavy weighting cannot usually occur in the case of "aunts," for they are, as a rule (owing to selection to the third generation) daughters of old and famous sires, with plenty of material for basing averages upon. We do not get *many* "nieces" attached to "aunts," who are not daughters of famous sires. Such is probably the source of divergence in σ_a between nieces and sisters.

Using formula (xxviii.), on p. 274, we find

$$r' = ·0114,$$

and for the regression coefficient ·0128.

This value is much below the ·05 of the law of ancestral heredity, and below the reduced value ·02, which we might have expected to reach. Still, it again shows

---

* Every dam appears as a granddam, otherwise the fecundity of the daughter could not have been found.

positive correlation, and we may conclude that the patent character in the daughter is inherited latently through the male line.*

But there is another and far more significant method of looking at this result, namely, by considering the meaning of R on p. 274. We may treat the fecundity of daughters as really a character of the sire, and their mean fecundity as a measure of a latent character in him. R is then the correlation between a latent character in both a stallion and his sire, and we see that it is sensibly inherited for $R = \cdot 1174$. To compare with the law of ancestral heredity, we must use the coefficient of regression, for the stallions are much more variable than their sires. We find

$$\text{Regression of stallion on sire} = \cdot 2056,$$

which carries us a long way in the direction indicated by that law. Thus it is extremely probable that this law of inheritance applies not only to the inheritance of a patent character, or of a character latent in one sex and transmitted to a second, but also to the inheritance of a character latent both in the transmitter and receiver. The present method accordingly seems applicable to the inheritance of a character latent in two individuals, if we take the mean of the character, when patent in the offspring, as a measure of its strength in the individual in whom it is latent. If l_1 be the measure of a latent character in a parent, then the offspring will have a mean value $ql_1 + c_1$ of this character, where q is the coefficient of parental regression and c_1 a constant. If l_2 be the measure of the same latent character in a relative, then the offspring in this case will have $ql_2 + c_2$ of the character. But the correlation of l_1 and l_2 will be identical with that of $ql_1 + c_1$ and $ql_2 + c_2$, as I have shown elsewhere.† Thus the mean of the patent character in the offspring may be used to measure the correlation between latent characters in their parents.

To sum up our results for thoroughbred mares, we conclude that their fecundity, notwithstanding the imperfections and difficulties of the record, has been demonstrated to be inherited, and this, both through the male and female line, so far as we can judge, with an equal intensity. The *apparent* value of this intensity, except in the case of latent characters, is much below that required by the law of ancestral heredity, roughly, perhaps, 2/5 of that value; but there is considerable reason to think that this reduction may take place owing to the presence of fictitious values in the record arising from the peculiar circumstances under which thoroughbred horses are reared and bred. These fictitious values would hardly influence the means and variability of arrays like they must do the relationship between pairs of individuals. Hence, when we deal with such means and variabilities as in the cases on pp. 309 and 313, we find a much closer approach to the law of ancestral heredity. Fecundity is certainly inherited; that it is inherited according to the Galtonian law

---

* As a matter of fact, this conclusion is stronger than it appears here, for the correlation between nieces and aunts was worked out, without grouping, for fourteen distinct series, and in *thirteen* of them was found to be sensibly positive; in the fourteenth it was found to have an insignificant negative value.

† "On the Reconstruction of the Stature of Prehistoric Races," 'Phil. Trans.,' A, vol. 192, p. 183.

is not demonstrated, but may be treated as probable until the results of further investigations—preferably by breeding experiments instituted for this very purpose—are available.

(18.) *Conclusion.*—The investigations of this memoir have been to some extent obscure and difficult, but the general result is beyond question.

Fertility and fecundity, as shown by investigations on mankind and on the thoroughbred horse, are inherited characters.

The laws of inheritance of these characters are with considerable probability those already developed in my memoir on the Law of Ancestral Heredity for the inheritance of directly measurable organic characters.

In the course of the work it has been shown how a numerical measure may be obtained for the inheritance of a character by one sex from the other, when it is patent in the former and latent in the latter. Fertility and fecundity purely latent in the male (in the sense here used) are shown to be transferred by him from his mother to his daughter. Thus DARWIN'S views with regard to the transmission through one sex of a character peculiar to the other are given a quantitative corroboration.*

When we turn from these points to their weight and importance for the theory of evolution, we are at once encountered by all the wide-reaching principles which flow from the demonstration that genetic (reproductive) selection is a true factor of development. Let us look at these a little more closely.

If natural selection were to be absolutely suspended, *i.e.*, if there were no differential death-rate at all, then development would not for a moment cease. Not only is fertility inherited, but there can be small doubt that it is closely correlated with all sorts of organic characters; thus the inheritance of fertility marks, the moment natural selection is suspended, a progressive change in a great variety of organic characters. Without a differential death-rate the most fertile will form in every generation a larger and larger percentage of the whole population. There are very few characters which may not be supposed to be more or less directly correlated with fertility, and in reproductive selection we see a cause of progressive change continuously at work.† There is, so to speak, in every species an innate tendency to progressive change, quantitatively measurable by determining the correlation coefficients between fertility and organic characters, and between fertility in the parents and in the offspring. This "innate tendency" is no mysterious "force" causing evolution to take place in a pre-ordained direction; it is simply a part of the physical organisation of the individual, which does not leave fertility independent of

---

* The method is perfectly general, and a value can always be found for the intensity of transmission of a sexual character through the opposite sex. We could obtain, for example, a numerical measure of the manner in which a bull transmits good milking qualities to its offspring.

† I have endeavoured to show ('Roy. Soc. Proc.,' vol. 59, p. 303), that fertility is correlated with stature in woman. I hope later to return to the correlation of fertility and physique.

physique and organic relationship, or leave these characters uncontrolled by the principle of heredity. It seems to me, therefore, that the results of this memoir force on us some modification of current views of evolution. The suspension of natural selection does not denote either the regression of a race to past types, as the supporters of panmixia suggest, or the permanence of the existing type, as others have believed. It really denotes full play to genetic or reproductive selection, which will progressively develop the race in a manner which can be quantitatively predicted when once we know the numerical constants which define the characters of a race and their relation to racial fertility. In other words, natural selection must not be looked upon as moulding an otherwise permanent or stable type; it is occupied with checking, guiding, and otherwise controlling a progressive tendency to change.

So soon as a species is placed under a novel environment, either artificially or naturally, the equilibrium is disturbed, and it will begin to progress in the manner indicated by genetic (reproductive) selection, until this progress is checked by the development of characters in a manner or to an extent which is inconsistent with fitness to survive in the new surroundings. Within a very few generations a novel environment, sympathetic so to speak to the progressive tendency indicated by reproductive selection, produces the suitable variations without the assistance of natural selection. It seems to me that this principle ought to be borne in mind when, in laboratory experiments or in artificial breeding, natural selection is wholly or largely suspended, or again is altered in type; the species dealt with is unlikely to remain constant for several generations, but will develop in the direction indicated by genetic selection. Further, when stable types of life like the English sparrow are taken to America, or the English rabbit to Australia, where initially they fill a more or less vacant field among living forms, and natural selection is in part suspended, we should expect in a few generations a considerable divergence in type.* The converse aspect of the problem is also of great importance; namely, the natural selection of physical characters must tend to indirectly modify fertility and fecundity, if these be correlated with those characters. Variations in the fertility of local races need not be looked upon as due directly to environment, but may arise from the selection of characters correlated with fertility, combined with the law that fertility is itself an inherited character.

Lastly, the inheritance of fertility involves the "acceleration" of fertility; a race, natural selection being suspended, tends not only to increase but to increase at an increasing rate. This principle is again full of meaning, not only for the study of the manner in which lower types of life rapidly expand under changed environment, but also for the problems set to those philosophers who may desire that the most social and not the most fertile type of citizen may predominate in our modern civilised communities, where the state and public opinion to a greater or less extent hinder natural selection from playing the great part it does in wild life.

---

* It would be interesting to know whether the size or frequency of the litter of the Australian rabbit is greater than that of the English.

TABLE I.—Correlation of Fertility of a Woman and of her Daughter. Marriages completed, or having lasted at least 15 years in Daughter's case only. 4418 cases. Fertility of Mother weighted. (See p. 281.)

		Mother's fertility.																			Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Daughter's fertility.	0	28	42	70	98	109	90	71	81	67	52	29	22	3	3	4	3	1	..	2	775
	1	22	56	78	104	80	70	64	51	39	31	20	19	1	4	4	4	..	..	..	647
	2	21	39	63	65	78	73	47	47	40	29	20	11	2	2	..	2	..	..	2	541
	3	20	35	44	53	53	75	57	57	33	37	21	13	10	5	1	2	..	..	..	516
	4	17	27	58	52	69	62	53	46	41	24	16	12	11	6	1	..	..	..	..	495
	5	16	28	26	37	52	55	43	32	40	25	12	14	5	2	1	..	..	..	..	388
	6	13	20	26	42	41	39	30	38	28	19	5	8	3	2	1	..	..	..	1	316
	7	3	14	33	23	33	31	24	32	26	14	10	9	5	1	1	1	..	..	..	260
	8	6	9	15	16	25	16	22	15	20	11	6	6	6	..	1	3	..	..	..	177
	9	4	7	13	13	13	10	14	12	10	8	5	8	3	1	..	1	..	..	..	122
	10	..	13	2	8	10	7	13	7	8	6	1	3	1	1	..	..	..	..	..	80
	11	7	2	6	9	7	1	4	4	4	3	2	1	1	1	..	..	..	..	..	52
	12	..	1	2	1	4	4	..	2	1	..	2	1	2	..	..	..	..	..	..	20
	13	..	..	2	1	1	2	4	2	1	2	1	1	1	..	..	..	..	..	..	18
	14	..	1	..	1	..	2	1	..	..	1	..	..	..	..	..	..	..	..	..	6
	15	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	1
	16	1	2	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	4
Totals ..		158	296	439	523	575	537	447	426	358	263	150	128	54	28	14	16	1	0	5	4418

TABLE II.—Correlation between the Fertility of a Woman and that of her Daughter. No weight to the Fertility of the Mother. Marriage completed or having lasted at least 15 years in the Daughter's case only. 1000 cases from the Peerage. (See p. 282.)

		Mother's fertility.															Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Daughter's fertility.	0	3	6	18	15	15	12	11	5	6	3	1	1	..	..	1	97
	1	6	13	16	23	17	16	10	11	6	7	2	..	..	..	..	127
	2	6	7	17	22	17	15	9	14	7	4	2	..	..	..	..	120
	3	8	12	18	20	13	22	12	12	10	6	1	3	..	..	..	137
	4	5	6	16	15	21	19	22	8	12	5	3	4	1	..	..	137
	5	4	7	4	12	16	16	17	11	13	7	6	3	..	..	..	116
	6	1	9	8	12	13	9	13	9	6	6	2	4	1	..	..	93
	7	1	2	5	6	9	9	5	7	5	6	2	2	2	..	1	62
	8	..	1	6	2	8	6	5	5	5	3	2	..	2	..	..	45
	9	1	1	1	5	4	4	6	1	1	1	2	..	..	..	..	27
	10	..	3	1	2	2	2	..	4	3	2	..	1	..	1	..	21
	11	..	..	1	2	1	1	2	..	..	..	1	1	..	..	..	9
	12	..	..	..	..	2	..	..	..	..	..	..	..	2	..	..	4
	13	..	..	..	..	..	1	2	..	..	..	..	..	1	..	1	5
Totals ..		35	67	111	136	138	132	114	87	74	50	24	19	9	1	3	1000

TABLE III.—Correlation between the Fertility of a Woman and that of her Daughter. No weight to Fertility of the Mother. Marriage completed or having lasted at least 15 years in Daughter's case only. 1000 cases, largely "Heiresses" from the Landed Gentry. (See p. 283.)

		Mother's fertility.																					Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Daughter's fertility.	0	20	12	9	6	18	16	14	15	7	8	3	5	1	..	..	..	..	..	..	.	..	134
	1	16	10	17	23	12	11	11	7	3	7	4	..	..	..	..	..	..	..	..	..	..	121
	2	6	11	11	9	18	12	18	9	2	2	1	2	..	1	..	..	..	..	..	..	..	102
	3	15	9	8	13	19	10	6	10	4	4	1	..	..	1	..	..	..	..	..	..	..	100
	4	23	8	7	17	11	9	9	11	7	3	2	3	2	..	..	..	..	..	..	..	..	112
	5	17	8	8	16	13	10	12	10	6	6	2	..	3	..	..	..	..	..	..	..	..	111
	6	12	3	2	15	9	8	8	7	3	4	5	1	..	2	1	..	..	..	..	..	..	80
	7	14	3	11	3	7	7	7	9	4	3	3	..	3	1	1	..	..	..	..	..	..	76
	8	9	5	4	6	6	6	5	2	3	2	2	2	..	1	..	2	..	..	..	..	..	55
	9	10	3	1	2	1	2	5	6	..	1	1	2	1	2	..	..	..	..	..	..	..	37
	10	2	..	3	3	4	2	5	6	1	2	..	2	..	..	1	..	..	..	..	..	..	31
	11	4	..	2	..	3	2	1	3	1	..	..	..	..	..	..	..	..	..	..	..	..	16
	12	1	1	1	1	2	1	..	..	2	1	1	1	1	..	..	..	..	..	..	..	..	13
	13	..	..	..	..	2	..	1	2	..	..	..	..	1	..	..	..	..	..	..	..	1	7
	14	..	..	..	..	1	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
	15	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	1
	16	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	17	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	18	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	19	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	21	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
	22	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..		149	73	84	114	126	98	103	97	43	43	26	18	12	8	3	2	0	0	0	0	1	1000

TABLE IV.—Correlation between the Fertility of a Woman and that of her Mother, when the Marriages of both have lasted at least 15 years and no weight is given to the Fertility of the Mother. 1000 cases from the Peerage. (See p. 285.)

Daughter's fertility.	Mother's fertility.																	Totals.
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
0	:	5	9	11	18	21	15	8	9	6	3	2	3	:	:	:	:	110
1	:	12	5	14	15	10	13	9	8	5	3	2	2	:	:	:	:	98
2	:	9	9	10	15	18	15	9	3	2	4	2	:	:	:	:	1	97
3	:	5	10	16	11	9	14	13	10	4	8	2	3	:	:	:	:	105
4	:	5	5	19	17	21	15	18	10	14	2	1	5	1	:	:	:	133
5	:	7	6	7	17	23	9	12	13	14	8	3	2	2	:	:	:	123
6	:	4	5	8	11	15	12	15	14	7	5	3	3	1	:	:	:	103
7	:	5	4	3	8	4	13	9	8	5	10	2	1	1	:	:	:	73
8	:	1	2	4	12	9	9	8	5	12	3	4	1	2	1	:	:	73
9	:	:	:	4	3	3	4	7	5	3	2	2	1	:	:	:	:	34
10	:	:	:	1	2	1	3	4	6	3	2	:	1	:	1	:	:	24
11	:	:	:	2	1	1	1	:	:	1	2	:	:	:	:	:	:	8
12	:	:	2	1	2	3	:	1	1	:	:	1	:	1	:	1	:	13
13	:	:	:	:	:	2	1	:	:	:	:	1	:	2	:	:	:	6
14	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	0
15	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	0
16	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	0
Totals . .	0	53	57	100	132	140	124	113	92	76	52	25	22	10	2	1	1	1000

TABLE V.—Correlation of the Fertility of a Man and of his Son as shown in Marriage. The Marriage of the Son being completed or having lasted at least 15 years. Weight given to the Fertility of Father. 6070 cases. (See p. 287.)

Number of son's children.	Number of father's children.																								Totals.
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
0	..	36	67	79	78	97	86	70	72	36	35	19	8	8	5	10	2	..	..	1	..	..	1	..	710
1	..	62	89	114	104	97	77	81	69	50	25	24	11	7	1	6	4	..	1	..	..	..	..	1	823
2	..	58	75	87	106	103	85	75	68	41	33	23	16	15	2	2	5	..	..	..	..	1	..	..	795
3	..	39	84	89	98	87	87	83	65	54	40	28	10	8	7	6	1	1	1	2	..	..	..	..	790
4	..	38	66	84	90	89	76	73	68	51	27	22	15	14	5	2	3	1	..	1	..	..	..	..	725
5	..	37	53	70	81	68	58	66	61	41	31	29	17	5	3	4	1	..	..	..	..	..	..	..	625
6	..	22	43	55	59	67	41	41	49	40	17	22	9	5	1	..	3	1	..	..	..	1	..	..	476
7	..	18	34	35	30	34	51	39	42	33	24	15	5	5	1	2	..	..	..	1	..	..	..	..	369
8	..	18	20	26	36	36	28	30	25	17	12	12	7	3	1	2	2	..	..	1	..	..	..	..	276
9	..	10	12	22	24	27	21	24	14	11	9	4	5	1	2	..	1	..	..	2	..	..	1	..	190
10	..	4	11	17	13	18	14	10	12	14	6	4	2	2	..	..	..	..	..	..	..	..	..	..	127
11	..	4	7	6	7	18	7	3	3	3	6	3	3	..	1	1	..	..	..	1	..	..	..	..	73
12	..	4	..	3	5	6	2	2	3	1	5	2	1	..	1	..	..	..	..	..	..	..	..	..	35
13	..	..	..	3	1	4	2	1	3	2	..	3	..	..	..	..	..	..	..	..	..	..	..	..	19
14	..	..	1	..	2	2	2	4	..	..	1	1	..	..	..	..	..	1	..	..	..	..	..	..	14
15	..	..	..	..	2	..	1	..	1	..	..	2	..	..	2	..	..	..	..	..	..	..	..	..	8
16	..	..	1	..	1	..	..	1	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	5
17	..	..	..	..	..	1	1	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	3
18	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	1
19	..	..	..	..	..	..	..	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	2
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
21	..	..	..	..	1	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
22	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
23	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
24	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..	0	351	563	690	739	754	640	604	557	394	272	214	110	73	32	35	22	4	2	9	0	2	2	1	6070

TABLE VI.—Correlation of the Fertility in Marriage of a Man and his Son; no weight being given to the Fertility of the Father, and the Marriage of the Son being completed or having lasted at least 15 years. 1000 cases from the Peerage. (See p. 287.)

		Fertility of the father.																	Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Fertility of sons.	0	5	8	7	14	18	2	2	3	8	3	4	4	..	..	..	..	..	78
	1	3	3	6	5	8	8	6	5	4	..	2	..	1	..	..	..	..	51
	2	7	5	6	13	12	12	12	6	5	4	2	1	1	..	..	..	..	86
	3	5	10	13	11	17	13	13	12	10	4	1	2	1	..	..	..	..	112
	4	4	16	18	24	23	5	18	10	7	8	1	5	2	1	..	..	1	143
	5	9	8	11	14	16	12	16	12	2	5	8	6	3	..	2	..	..	124
	6	3	4	10	16	13	11	11	10	10	1	2	2	1	..	..	..	..	94
	7	5	6	8	7	10	14	11	12	4	7	1	1	..	..	..	..	..	86
	8	3	5	4	15	19	7	10	8	4	2	2	1	..	..	1	..	..	81
	9	1	6	5	9	5	5	8	5	4	3	3	1	..	..	..	..	..	55
	10	2	3	9	3	2	5	4	6	4	3	1	1	1	..	..	..	..	44
	11	1	1	1	4	2	1	1	2	1	1	..	..	..	..	..	..	..	15
	12	..	..	2	2	2	..	1	1	1	2	..	1	..	..	..	..	..	13
	13	1	..	..	1	3	..	1	1	1	1	..	1	..	1	..	..	..	10
	14	..	1	..	..	..	1	..	..	..	..	1	..	..	..	..	..	..	3
	15	..	..	1	..	..	..	..	1	..	..	1	..	..	..	..	..	..	3
	16	..	..	..	..	..	..	..	..	1	1	..	..	..	..	..	..	..	2
Totals ..		49	76	101	138	150	96	114	94	66	45	29	26	10	2	3	0	1	1000

TABLE VII.—Correlation of the Fertility in Marriage of a Man and his Son; no weight being given to the Fertility of the Father, and the Marriages of both Father and Son having lasted at least 15 years. 1000 cases from the Landed Gentry. (See p. 288.)

		Fertiliy of father.																Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Fertility of the son.	0	1	1	..	11	9	6	4	5	6	2	1	..	2	..	..	..	48
	1	..	5	7	11	5	8	7	2	7	1	..	2	..	1	..	..	56
	2	2	6	9	12	8	7	8	4	1	4	3	1	1	..	..	..	66
	3	5	8	15	14	11	10	10	10	7	5	3	4	..	..	..	..	103
	4	4	9	14	22	16	21	10	12	11	7	1	8	..	2	1	..	138
	5	5	7	12	19	30	12	17	13	12	10	3	3	2	..	..	..	145
	6	5	4	9	12	17	16	20	14	8	8	5	4	..	3	..	..	125
	7	..	4	5	23	13	12	16	8	9	6	3	2	..	2	1	..	104
	8	3	2	5	7	8	12	7	11	8	5	4	3	..	1	..	1	77
	9	1	..	6	8	3	7	5	6	3	1	2	1	..	..	..	1	44
	10	..	1	5	7	4	3	4	7	1	4	2	1	..	1	1	..	41
	11	..	1	1	3	3	4	5	2	4	2	2	..	..	..	..	..	27
	12	..	..	1	..	2	1	..	4	..	3	1	..	..	2	..	..	14
	13	..	..	2	1	1	1	1	..	..	..	1	..	..	1	..	..	8
	14	..	..	..	1	1	..	..	..	..	..	..	..	..	..	..	..	2
	15	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	16	..	..	..	..	..	..	..	..	1	..	..	..	..	..	..	..	1
	17	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	1
Totals ..		26	48	91	151	132	120	114	98	78	58	31	29	5	13	3	3	1000

TABLE VIII.—Correlation of the Fertility of a Woman and her Paternal Grandmother. No weight being given to the Fertility of the Grandmother, and the Marriages of both having lasted for at least 15 years. 1000 cases from the Peerage and Baronetage. (See p. 289.)

		Fertility of paternal grandmother.																					Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Fertility of granddaughter.	0	6	12	14	11	15	15	8	5	6	2	2	2	1	1	..	..	..	..	..	..	1	101
	1	10	10	7	13	13	5	7	4	2	7	1	1	..	..	..	..	..	..	..	..	..	80
	2	6	7	20	14	11	10	11	6	6	4	2	1	..	..	..	1	..	..	..	..	..	99
	3	9	15	12	20	13	9	15	14	10	1	..	5	..	..	..	1	..	..	1	..	..	125
	4	7	13	16	19	19	11	11	13	8	2	5	4	..	..	..	1	..	..	..	..	..	129
	5	4	9	10	14	20	15	16	12	9	9	5	2	..	..	..	..	..	..	..	..	..	125
	6	5	10	17	9	16	10	8	8	6	5	5	1	2	1	..	..	..	..	..	..	..	103
	7	7	4	8	7	8	9	10	15	10	5	2	1	2	..	..	..	..	..	..	..	..	88
	8	6	7	5	5	3	6	9	12	6	1	2	1	..	..	..	1	..	..	..	..	..	64
	9	3	3	3	4	4	4	1	2	6	..	1	..	2	..	..	..	..	..	1	..	..	34
	10	2	1	4	2	3	1	1	6	1	5	1	..	..	..	..	..	..	..	..	..	..	27
	11	1	..	2	2	..	2	1	..	..	..	1	..	..	..	..	..	..	..	..	..	..	9
	12	..	..	..	2	..	5	1	1	1	..	1	1	..	..	..	..	..	..	..	..	..	12
	13	1	..	..	1	..	..	..	..	1	..	..	1	..	..	..	..	..	..	..	..	..	4
Totals ..		67	91	118	123	125	102	99	98	72	41	28	20	7	2	0	4	0	0	2	0	1	1000

TABLE IX.—Correlation between the Apparent Fertilities of Mares and Dams, with a minimum of at least four coverings. 1100 cases. (See p. 291.)

		Mare's foals.																			Totals.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Dam's foals.	1	..	..	..	..	..	1	..	1	..	..	1	..	..	2	..	..	..	..	..	5
	2	..	1	8	7	4	4	6	5	3	2	1	2	3	..	..	..	..	..	..	46
	3	..	2	5	5	3	5	10	7	5	3	4	1	2	3	..	1	..	..	..	56
	4	..	1	5	7	12	15	13	9	13	8	7	4	4	4	2	3	1	1	..	109
	5	..	2	6	10	7	13	18	7	11	8	6	10	2	1	3	1	..	1	..	106
	6	..	7	10	17	16	14	6	9	8	10	7	2	5	1	4	..	1	..	..	117
	7	1	2	15	17	22	18	10	10	14	8	6	2	2	..	1	..	1	..	1	130
	8	1	1	7	17	20	19	11	14	16	6	4	9	2	2	..	..	..	..	..	129
	9	1	1	8	15	6	6	7	8	6	1	6	3	2	2	1	..	..	..	..	73
	10	..	2	4	9	15	16	12	9	4	5	6	2	2	3	1	..	..	..	..	90
	11	..	2	7	14	11	7	10	13	8	4	..	3	2	..	1	..	..	..	..	82
	12	..	2	6	7	6	11	8	8	5	..	1	2	..	1	..	..	..	..	..	57
	13	..	..	2	4	9	3	3	5	3	4	1	..	..	..	..	..	..	..	..	34
	14	..	1	4	3	6	5	5	5	2	2	4	..	..	..	..	..	..	..	..	37
	15	..	2	..	1	1	1	4	1	1	3	..	..	1	..	1	..	..	..	..	16
	16	..	..	..	..	..	..	1	1	3	..	..	1	..	..	1	..	..	..	..	7
	17	..	..	..	..	..	2	..	1	1	1	..	1	..	..	..	..	..	..	..	6
	18	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
Totals ..		3	26	27	133	138	140	124	113	103	65	54	42	27	19	15	5	3	2	1	1100

TABLE X.—Correlation between the Apparent Fertility of Mares and Dams, with a minimum of at least eight coverings. 1100 cases. (See p. 292.)

		Mare's foals.																		Totals.
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Dam's foals.	1	..	..	..	..	1	..	..	..	..	..	..	..	1	..	..	..	..	..	2
	2	..	..	1	1	2	..	2	1	1	..	..	..	..	..	..	..	..	..	8
	3	1	..	1	..	1	2	4	..	2	..	2	..	2	..	..	1	..	..	16
	4	..	1	1	1	7	7	6	10	5	3	8	2	5	..	2	1	1	..	60
	5	..	1	1	5	7	15	9	13	12	1	11	4	3	3	..	2	1	..	88
	6	..	1	3	9	17	16	14	15	12	16	5	5	4	5	..	1	..	..	123
	7	2	2	7	16	18	17	24	22	19	9	9	1	3	3	1	1	..	1	155
	8	2	1	6	6	18	21	25	26	15	7	15	5	6	2	1	1	..	..	157
	9	..	1	6	4	9	10	20	9	4	9	7	3	3	5	..	..	..	..	90
	10	..	..	3	8	14	17	16	8	9	9	7	2	4	1	1	..	..	..	99
	11	..	1	3	4	10	19	18	14	7	1	6	2	1	1	..	..	..	..	87
	12	..	1	2	6	11	11	14	12	6	6	4	1	1	1	..	..	..	..	76
	13	..	..	1	4	7	11	8	9	9	1	3	1	..	..	..	..	..	..	54
	14	..	2	2	1	4	7	6	3	3	4	2	..	..	..	..	..	..	..	34
	15	..	..	..	1	2	5	2	4	6	1	..	2	..	1	..	..	..	..	24
	16	..	..	..	1	..	3	2	6	..	..	1	..	..	1	..	..	..	..	14
	17	..	..	..	..	2	..	2	1	1	..	2	..	1	..	..	..	..	..	9
	18	..	..	..	..	..	2	..	1	..	..	..	..	..	1	..	..	..	..	4
Totals..	..	5	11	37	67	130	163	172	154	111	67	82	28	34	24	5	7	2	1	1100

Table XI.—Correlation between the Fecundity and the *Apparent* Fertility of Brood-mares. 1000 cases. (See p. 300.)

		Apparent fertility.																			Totals.
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Fecundity.	*a*	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	2
	b	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
	c	..	2·5	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	3·5
	d	..	5·5	2	..	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	8·5
	e	..	13	5	3·5	2	..	1	..	..	..	..	..	..	..	..	..	..	..	..	24·5
	f	..	..	10	4·5	9	4	3	..	..	..	..	..	..	..	..	..	..	..	..	30·5
	g	..	..	16	13	6	13	9	5	1	1	..	..	..	..	..	..	..	..	..	64
	h	..	..	17·5	5	13	12·5	11	10	6·5	5·5	1·5	..	..	..	..	..	..	..	..	82·5
	i	..	..	16·5	5	6	19·5	23	14	17·5	13·5	6·5	3	1	..	..	..	..	..	..	125·5
	j	..	..	..	21	21	21	11	17	18	4	21	6	2	3	..	..	..	..	..	145
	k	..	..	4	..	22	..	9	21·5	12	18	5	15	9	2	3·5	..	..	..	..	121
	l	..	1	1	45	..	13	19	10·5	11	12	6	5	8	9	3·5	2	2	..	..	148
	m	..	..	..	..	31	13	..	10	17	10	12·5	7	3	4	6	3·5	..	1	..	118
	n	..	..	..	..	..	13	9	10	10	1·5	3·5	2	3	..	1	2·5	2	1	·5	59
	p	..	..	..	..	..	..	..	..	..	1·5	4	5	2	2	1	..	1	1	·5	18
	q	..	..	..	..	14	9	5	10	4	2	..	..	2	2	1	..	..	..	..	49
Totals ..		2	23	73	97	125	118	100	108	97	69	60	43	30	22	16	8	5	3	1	1000

TABLE XII.—Correlation Table for 1200 Mares and their Dams. Fecundity with a minimum of four coverings. (See p. 303.)

	Mares.																Totals.
Dams.	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
a	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
b	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	1
c	..	..	..	..	..	·5	..	..	1	·5	..	..	..	..	..	..	2
d	..	..	..	..	..	·5	..	1	1	2·5	..	3	1	..	..	..	9
e	..	..	..	..	2·5	1	1·5	1	3	2·5	..	2·5	1	·5	·5	·5	16·5
f	..	..	1	..	·5	1	4·5	2·5	2·5	·5	2	3·5	3	1	1	1·5	24·5
g	..	..	..	1	2	3	5	2·5	10·5	10	3	6·5	1	1	1	4	50·5
h	·5	..	·5	·5	3	5·5	9	10·25	10·25	14·5	14·25	15·75	13·5	4	3	6·5	111
i	·5	..	..	1	4·5	3	12	14·25	19·25	21·5	17·25	18·25	10	5·5	1·5	4·5	133
j	1	..	..	5	3	3	10	11	18·5	24	34	26	21·5	8	3·5	12	180·5
k	..	..	1	1·5	5·5	4	14	17	21	19	28·75	31·75	23	13·25	3·75	8·5	192
l	1	1	1·5	3	6·5	7	11	13·5	17·5	32	25·75	27·75	15·5	11·25	2·25	9·5	186
m	..	..	·5	3	3	7·5	5·5	9·25	13·75	19·5	18·5	24	14	8	3	10	139·5
n	..	..	·5	..	1·5	4·5	5·5	7	6·5	10·5	9·5	13	11·5	9·5	1	7·5	88
p	..	..	..	..	1	·5	..	1·25	2·25	6	2·5	5·5	4	..	..	2·5	25·5
q	..	..	.	1	1	..	4	3·5	4·5	9	5	4	3·5	2·5	1	2	41
Totals ..	3	1	5	16	34	41	82	94	131·5	172	160·5	182·5	122·5	64·5	21·5	69	1200

TABLE XIII.—Correlation between Brood-mares and their Dams with regard to Fecundity with a minimum of eight coverings. 1000 cases. (See p. 304.)

		Fecundity of mares.																Totals.
		a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
Fecundity of dams.	*a*	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	b	..	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	1
	c	..	..	..	..	..	·5	..	..	1	·5	..	..	..	..	..	..	2
	d	..	..	..	..	..	·5	..	..	1	·5	..	1	..	..	..	..	3
	e	..	..	..	..	..	1	1	1	..	1	·5	2	..	..	..	..	6·5
	f	..	..	1	..	..	1	4	3·5	2·5	1	3·5	2	2	..	1	..	21·5
	g	..	..	..	..	1·5	2·5	3	2	11	10	7·5	4·5	4	4	2	..	52
	h	..	..	·5	..	2·5	5	8	10	9·5	17	18·25	10·25	10·5	5	3·5	..	100
	i	..	..	1·5	2	2	2·5	9	12·5	22	22	24·75	16·75	15·5	7	2·5	2	142
	j	..	..	..	1	1·5	4·5	9	8	17	25	34	22	20·5	6	4·5	4	157
	k	..	..	·5	·5	2·5	3·5	5·5	14·25	22·25	19·5	31·5	31	19	9·5	6	2	167·5
	l	..	1	·5	1·5	5	4	5	14·25	19·25	25·5	26	20	17	5·5	3	2	149·5
	m	..	..	·5	3·5	..	2	3	9·25	13·25	13·5	12	17·5	8·5	5·25	3·75	1	93
	n	..	..	1	..	..	6	5	6·5	8	13	8·5	12	11·5	3·75	2·25	2	79·5
	p	..	..	..	..	..	·5	..	2·25	2·25	7·5	2	4	2·5	1·5	..	..	20·5
	q	..	..	..	..	..	..	..	·5	·5	2	1	..	..	..	..	1	5
Totals ..		0	1	5·5	8·5	15	33·5	52·5	82	129·5	158	169·5	144	111	47·5	28·5	14	1000

TABLE XIV.—Correlation of the Fecundity of Brood-mares and their Dams with a minimum of four coverings. Second series. 1200 cases. (See p. 305.)

		Fecundity of mares.																Totals.
		a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
Fecundity of dams.	*a*	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	b	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	c	..	..	..	..	..	..	..	1·25	·25	..	1	..	..	..	..	..	2·5
	d	..	..	..	..	..	1	..	·75	·75	..	..	..	..	..	..	..	2·5
	e	..	..	..	..	..	1	1	·25	·25	..	·5	1·5	·5	1·5	..	..	6·5
	f	..	..	..	..	1	..	5	3·25	2·25	5	2·5	1·5	1·5	·5	1	..	23·5
	g	..	..	..	..	2·5	1·5	6	5	8	9	7	7·5	6	7	1	..	60·5
	h	·5	..	1·25	1·25	4	3	6	8·25	9·75	19	11	9·5	13·25	7·5	·25	3	97·5
	i	1·5	..	2·25	3·25	3	3	6	9·25	14·75	30	15	15·5	16·75	6	2·25	5	133·5
	j	..	..	·5	3·5	3·5	4·5	16	12·5	20·5	27	27·5	28·5	28	11	2	6	191
	k	2	..	..	·5	5	4	13	18·25	28·75	23·5	20·25	25·75	23	12·75	2·75	13·5	193
	l	1	..	..	2·5	6	3	4	22·25	23·75	24	27·75	18·25	22·5	13·25	4·75	13·5	186·5
	m	..	..	·5	2	4·5	5	10	7·75	10·25	18	15	31	12	9·5	2	5·5	133
	n	1	..	1·5	1	2·5	5	7	5·5	7	14·5	11·5	10·5	9·25	4·75	1·5	8	90·5
	p	..	..	..	1	1	..	..	·25	1·25	6	2·5	5·5	3·25	3·25	..	2·5	26·5
	q	..	..	..	..	..	1	2	2·5	5·5	6	9	10	7·5	2	2·5	5	53
Totals ..		6	0	6	15	33	32	76	97	133	182	150·5	165	143·5	79	20	62	1200

TABLE XV.—Correlation of the Fecundity of Brood-mares and of their Granddams, with a minimum of four coverings. 1000 cases. (See p. 305.)

	Fecundity of mares.																Totals.
Fecundity of granddams.	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*p*	*q*	
a	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	..	..	..	..	..	..	..	..	..	..	1	..	..	..	..	..	1
c	..	..	..	..	·5	..	·5	·75	1·25	·5	..	1	·5	..	..	..	5
d	..	..	..	..	·5	..	·5	·25	·75	1·5	1	1	·5	..	..	..	6
e	..	1	..	..	..	·5	..	·5	1·5	1·5	..	3	..	·5	..	..	8·5
	..	..	..	..	..	1·5	..	2·5	3·5	5·5	4·5	3·5	6·5	2	..	1	30·5
g	1	..	..	4	3	4	4	5	7	7	7·5	15·5	12·5	3·5	..	6	80
h	..	..	1	1·5	4·5	4	6	7·5	12·5	11	11·5	19	10·25	5·25	..	5·5	99·5
i	..	..	2	2·5	7	5·5	6	13·5	14·5	18·5	15·5	29	13·75	8·75	1	8·5	146
j	1	..	2	3	7	6	8	10·5	14·5	16	18	22	19	11·5	·5	17	156
k	1	..	1	3	4	7	11·5	11·25	15·25	25	10·75	21·75	10·5	14	1	13	150
l	..	..	..	2	2	2	7·5	8·75	12·75	25·5	10·75	20·75	16	5	1·5	9	123·5
m	..	..	..	..	7	4	3·5	8	13	5	11	11	14	5·5	·5	6·5	89
n	1	..	..	·5	2	3	4·5	3·5	6·5	8·5	4·5	6	4	2·5	..	3·5	50
p	..	..	..	·5	..	..	1	1·5	1·5	1·5	·5	4	1·5	1	..	4	17
q	1	..	1	3	..	..	4	1	3	4	1·5	8·5	4·5	4·5	1	1	38
Totals ..	5	1	7	20	37·5	37·5	57	74·5	107·5	131	98	166	113·5	64	5·5	75	1000

TABLE XVI.—Correlation Table for Weighted Mean Fecundities of Arrays of Stallion's Daughters and Stallion's Sire's Daughters. 138,424 cases of Aunts and Nieces. (See p. 311.)

Stallion's sire.		Stallion.																											Totals.
		c		*d*		*e*		*f*		*g*		*h*		*i*		*j*		*k*		*l*		*m*		*n*		*p*		*q*	
		3–3·5.	3·5–4.	4–4·5.	4·5–5.	5–5·5.	5·5–6.	6–6·5.	6·5–7.	7–7·5.	7·5–8.	8–8·5.	8·5–9.	9–9·5.	9·5–10.	10–10·5.	10·5–11.	11–11·5.	11·5–12.	12–12·5.	12·5–13.	13–13·5.	13·5–14.	14–14·5.	14·5–15.	15–15·5.	15·5–16.	16–16·5.	
e	4·5–5	..	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	1
	5–5·5	..	..	..	..	..	..	..	..	..	..	·25	·25	..	..	..	..	1	1	..	..	..	·25	·25	..	..	..	..	3
f	5·5–6	..	..	..	..	..	..	..	..	..	..	..	..	..	60·25	·25	..	1	1	..	..	..	..	..	..	..	..	..	62·5
	6–6·5	..	..	..	..	..	..	..	..	..	..	·75	·75	..	·25	3·25	..	..	1·5	1·5	..	..	..	..	..	..	..	..	8
g	6·5–7	..	..	..	..	..	·5	·5	..	..	..	·75	·75	..	..	3	..	..	1·5	1·5	..	..	..	..	..	..	..	..	8·5
	7–7·5	..	..	..	..	..	·5	·5	..	..	..	..	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	5
h	7·5–8	..	..	..	..	..	..	..	..	..	..	28	24	..	4·25	16·25	..	..	..	..	..	..	..	..	..	..	2	2	76·5
	8–8·5	..	..	..	..	..	..	..	4	4	44·25	79·25	31	47·5	8·25	15·75	110	..	..	1·5	1·5	·25	·25	..	..	..	2	2	351·5
i	8·5–9	..	..	..	..	..	..	..	22	22	57·25	315·25	328	126·5	241·75	2162·75	1·5	396	5	30·5	163·5	18·25	·25	11	11	..	..	..	3912·5
	9–9·5	9	9	..	20	49	37·5	8·5	8·5	145	18	15	332·5	276	217·75	770·25	3603	317·5	196·5	226·5	22	31·5	9·5	..	..	..	20	20	6362·5
j	9·5–10	3	3	..	66	66	80·5	64·5	1·5	137	179·5	235·25	874·5	1415·25	3782·5	9020	5316	5197·5	1702·5	274·5	139	459·5	4	3	2	..	32	32	29,090·5
	10–10·5	..	..	..	107	107	74·5	56	127	221·5	428·5	357·75	1903	3554·25	3945·5	24,628	7165	7030·5	6263·5	640·5	152	202	161·5	5·5	42	..	114	114	57,400·5
k	10·5–11	..	3	3	19·5	40	243	10·5	22	22	76	224	705·5	676	2482·5	8148	2205·25	1502·75	1383	2122·5	406·75	72·25	470·5	14·5	..	..	73·5	73·5	20,999·5
	11–11·5	29·5	29·5	..	6·75	35·75	29·75	·75	30	18	301	67	245·5	747	526	944	1862·25	3692·25	3283·5	214	145·75	173·25	164	65·5	·5	·5	5·5	5·5	12,623
l	11·5–12	3·5	3·5	..	·75	6·25	13·25	7·75	23·75	23·75	44	99	86·25	367·25	1476·25	1347·75	1113·75	1877·25	343·5	263	72·5	20·5	4	..	7·5	7·5	10	10	7232·5
	12–12·5	..	..	..	..	..	..	..	2·25	2·25	..	7·5	25·75	·75	2·75	7·25	106·75	3·75	..	1	..	..	..	..	..	..	1	1	162
m	12·5–13	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1·5	1·5	3	3	..	32	..	..	..	..	..	..	41
	13–13·5	..	..	..	..	..	..	..	1·5	1·5	..	7·5	1·75	1	2·5	1·75	..	..	22·75	1·75	..	..	..	..	..	..	3	3	48
n	13·5–14	..	..	..	·25	·25	..	..	..	..	..	..	·25	1	1	·25	..	..	27·25	·25	..	..	..	·75	·75	..	..	..	32
	14–14·5	..	..	..	·25	·25	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·75	·75	..	..	..	2
p	14·5–15	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
	15–15·5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	0
q	15·5–16	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·5	·5	..	..	..	..	..	..	..	1
	16–16·5	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	·5	·5	..	..	..	..	..	..	..	1
Totals ..	..	45	48	3	220·5	304·5	479·5	149	242·5	597	1148·5	1437·5	4562	7214·5	12,751·5	47,068·5	21,485	20,021	13,235·5	3783	1104	1009·5	814·5	101·5	64·5	8	263	263	138,424